Hands-On Workshops at PharmaSUG 2004

CO-CHAIRS: Daphne Ewing, Synteract, Inc..
Jim Johnson, Covance, Inc.

Monday Morning, May 24th : Pearl Room / Third Level

9:00 AM-10:15 AM Paper Number: HW06 Audience Level: Advanced
Version 9 Changes and Enhancements
Dana Rafiee ~ Destiny Corporation

The primary goal of Version 9 is to provide support for a new level of computing that supports faster execution of applications, centralized access of data and support of the latest computing technology. These materials are designed as an overview of what is available and new and complement the online documentation that ships with the software.

Seating is limited and is on a first come, first served basis.

Room will be cleared between workshops.

10:30 AM-11:45 AM Paper Number: HW05 Audience Level: Intermediate
Programming Squared (Writing Programs that Write Programs)
Jim Johnson ~ Covance

This paper will demonstrate several techniques using Base SAS and macros that will take your programming skills to a new level by giving you the power to write programs that write programs. These simple techniques, which are powerful used alone, are awesome when combined. Using these techniques when your application is appropriate, you can write one program, which will act as though you wrote dozens of individual programs. The programming techniques will give you the ability to write more robust, highly flexible, intelligent programs that can extract data from little known sources, make decisions, write and execute program code depending less on user input. Programs using these techniques are much more difficult to debug, but once validated, can be used over and over by many projects. Although the techniques are simple, this paper is aimed at programmers who have the experience to know where and when to use these techniques efficiently.

Seating is limited and is on a first come, first served basis.

Hands-on Workshops (continued)

Monday Afternoon, May 24th : Pearl Room / Third Level

2:00 PM-3:15 PM Paper Number: HW08 Audience Level: Intermediate
An Animated Guide : The Data Step Debugger
Russell Lavery ~ Self

The Data Step Debugger (DSD) is shipped with every copy of base SAS but is rarely used. Running the DSD adds two windows to the normal SAS environment. These two new windows allow the programmer to step through his/her program – one line at a time- and watch the Program Data Vector (PDV) change. This ability to examine the PDV, as observations are being processed is a valuable debugging tool.

This presentation shows how to invoke the debugger, the common DSD commands, useful combinations of DSD commands and some tricks of the power user.

Seating is limited and is on a first come, first served basis.

Room will be cleared between workshops.

3:30 PM-4:45 PM Paper Number: HW07 Audience Level: Beginner
Effective Ways of Validating SAS Programs
Sy Truong ~ Meta-Xceed, Inc.

Validation is normally a laborious and arduous task. This paper will present methodologies and tools developed in SAS that will make the process painless. The goal is to add little or no effort from the user's perspective, yet gain the benefit of having a secured, audit trail of all SAS programs during the development and verification process. Some tasks and benefits covered in this paper include:

- preparing a requirement/functional specification and test plan
- comparing differences between different versions of programs
- adding notes describing edit changes to each version
- adding a validation checklist of tasks associated during verification and validation
- managing status of development to production by applying version numbers such as version 1.2
- generating reports for documentation and communication during validation
- summarizing test result reports

After you realize the ease of use and the amount of quality control that can be gained, the task of validation becomes routine and more transparent.

Seating is limited and is on a first come, first served basis.

Hands-on Workshops (continued)

Tuesday Morning, May 25th Workshops : Pearl Room / Third Level

9:00 AM-10:15 AM Paper Number: HW03 Audience Level: Beginner
READING AND WRITING XML FILES FROM SAS®
Miriam Cisternas ~ MGC Data Services / Ricardo Cisternas ~ MGC Data Services

XML (eXtensible Markup Language) is gaining popularity as a medium for data exchange. SAS Institute has added additional support recently for reading and writing XML files under both 8.2 and 9.1. The purpose of this workshop is to provide a quick introduction to XML, show the ways in which XML is read and written from SAS, and then present hands-on examples for reading in and mapping XML files to SAS datasets and writing out SAS datasets to XML files. Students will also gain experience using the newly released production version of XML Mapper; SAS's Java-based application to expedite the process of importing complicated XML documents into SAS datasets.

Seating is limited and is on a first come, first served basis.

Room will be cleared between workshops.

10:30 AM-11:45 AM Paper Number: HW01 Audience Level: Beginner
So You're Still Not Using PROC REPORT. Why Not?
Daphne Ewing ~ Synteract, Inc. / Ray Pass ~ Ray Pass Consulting

Everyone who can spell SAS knows how to use PROC PRINT. Its primary use may be as a development tool to help in debugging a long multi-step program, or as a simple report generator when all that is really needed is a quick look at the data, or even a basic low-level finished report. However, if a report generation/information delivery tool with powerful formatting, summarizing and analysis features is called for, then PROC REPORT is the solution. PROC REPORT can provide the standard PROC PRINT functionality, but in addition, can easily perform many of the tasks that you would otherwise have to use the SORT, MEANS, FREQ and TABULATE procedures to accomplish. PROC REPORT is part of the Base SAS product, can run in both an interactive screen-painting mode or a batch mode, and should be the basic tool of choice when there is a need to produce powerful and productive reports from SAS data sets. This paper will present the basics of PROC REPORT (non-interactive mode) through a series of progressively more sophisticated examples of code and output.

Seating is limited and is on a first come, first served basis.

Hands-on Workshops (continued)

Tuesday Afternoon, May 25th Workshops : Pearl Room / Third Level

2:00 PM-3:15 PM Paper Number: HW04 Audience Level: Intermediate
SAS with Style: Creating your own ODS Style Template for RTF Output
Lauren Haworth ~ Genentech, Inc.

Once you've started using the Output Delivery System, you'll quickly discover that your taste in output design probably doesn't coincide with the built in ODS styles shipped with SAS software. While you can edit your RTF output in Word to improve its appearance, a better approach is to create your own style template. This workshop will take you step by step through the process of creating a custom style for your RTF output.

You'll learn how to make minor modifications, and how to give your output a complete makeover. The workshop will also provide tips and tricks for taking advantage of the RTF destination, including the generation of custom page numbers and page breaks.

The workshop will walk through the TEMPLATE procedure, showing how you can redefine the default style elements and attributes to customize fonts, colors, and borders to suit your own personal or corporate style. You'll be given a basic style template that you can customize during the workshop and then take home to try out on your ODS output. The workshop is aimed at beginning to intermediate ODS users, and is based on SAS versions 8.2 and 9.

Seating is limited and is on a first come, first served basis.

Room will be cleared between workshops.

3:30 PM-4:45 PM Paper Number: HW02 Audience Level: Beginner
Five Ways to Create Macro Variables: A Short Introduction to the Macro Language
Art Carpenter ~ CA Occidental Consultants

The macro language is both powerful and flexible. With this power, however comes complexity, and this complexity often makes the language more difficult to learn and use. Fortunately one of the key elements of the macro language is its use of macro variables, and these are easy to learn and easy to use.

Macro variables can be created using a number of different techniques and statements. However the five most commonly methods are not only the most useful, but also among the easiest to master. Since macro variables are used in so many ways within the macro language, learning how they are created can also serve as an excellent introduction to the language itself. These methods include:

- %LET statement
- macro parameters (named and positional)
- iterative %DO statement
- using the INTO in PROC SQL
- using the CALL SYMPUT routine

Seating is limited and is on a first come, first served basis.

SAS at PharmaSUG 2004 (continued)

MONDAY, 2:00 PM-2:50 PM ~ Paper Number: SAS6
Room: Topaz Section: Tutorials

Next Step in Data Visualization: ODS Graphics Template Language

Written by Jeff Cartier.
Presented by Sanjay Matange, Senior Software Manager.

Now you can take the next step in data visualization by using Proc Template and ODS Graphics Template Language (GTL) to define your own custom graphs. Starting with syntax for defining familiar, simple charts, ODS GTL also provides powerful syntax to create your own custom layouts and overlays. Dynamic and macro variables can be stored in compiled graphical templates. These variables can be assigned values at run time to modify the output. Once a graphical template is stored, you can use the Data step to define or read a data set that serves as input to the template. ODS statements further define the output destination and image format for the graphical output.

In SAS 9.1, selected SAS/STAT and SAS/ETS procedures are able to produce sophisticated graphical output using this template language.

TUESDAY, 11:00 AM-11:50 AM ~ Paper Number: SAS1
Room: Diamond I Section: Applications Development

A Guide to Understanding Java Web Application Development

Written by Robert Girardin and Corey Benson.
Presented by Robert Girardin, Systems Developer.

You have been asked by your manager to port a legacy, fat-client application to the web. There are so many directions you can take in completing this project. It is hard to know where to start and what technologies to use.

The main focus of this paper is to introduce you to server-side Java. We will answer questions like:

- What is server-side Java technology?
- What are servlets, JavaServer Pages and custom tags?
- What are the benefits of using server-side Java technology?
- What is the Model-View-Controller(MVC) architecture?
- What is the Struts framework?
- What are the benefits of using webAF for Web application development?
- How do you customize the look of your Web application?

These topics will be covered while also highlighting how SAS AppDev StudioTM 3.0 makes it easier for you to build Web applications that are lightweight, easy to manage and instantly connect to SAS software.

TUESDAY, 1:30 PM-2:20 PM ~ Paper Number: SAS5
Room: Diamond II Section: Technical Techniques

Il n'est pas magique! Proc Report Compute Blocks Revealed

Written and presented by Sandy McNeill, Strategic Advisor for R&D.

If you have worked at all with Proc Report, you will probably have at one time or another had the occasion to use compute blocks. Were you frustrated when you didn't get the answer you thought you should? Frustrated about Missing Values? Compute Blocks are undeniably the power behind Proc Report. This paper will explain the internal workings of Proc Report Compute Blocks so that you'll never be mystified by them again.

SAS at PharmaSUG 2004 (continued)

TUESDAY, 2:00 PM-2:50 PM ~ Paper Number: SAS4
Room: Diamond I Section: Statistics and Pharmacokinetics

Progress in SAS Scientific Discovery Solutions: Genetics, Microarrays, and Proteomics

Written by Russ Wolfinger, Wendy Czika, and Kristen Kleiss.
Presented by Kristen Kleiss, Project Manager, Genomics R&D.

The SAS Genomics team is making steady progress in leveraging existing and new SAS technologies for more effective approaches to scientific discovery. In this presentation, we will overview enhancements to SAS/Genetics, SAS Research Data Management (RDM), and the SAS Microarray Solution, as well as introduce the two latest additions to the SAS Scientific Discovery Solutions line, SAS Genetic Marker and SAS Proteomics. Release 1.3 of the RDM platform contains several major improvements, including an expanded architecture for supporting input engines and analytical processes, as well as advanced GUI support via XML. SAS Genetic Marker is built on the RDM platform and features analytical processes invoking SAS/Genetics procedures in the background. SAS Proteomics is also an add-on to the platform and focuses on the management and analysis of spectral data, enabling advanced data analyses and profiling of peptides or metabolites. A live demonstration will highlight the basic features of the new solutions and how they enable robust identification of discriminating characteristics that can potentially be linked to life-changing diagnostic and/or therapeutic biomarkers.

TUESDAY, 2:00 PM-2:50 PM ~ Paper Number: SAS7
Room: Topaz Section: FDA Compliance: Electronic Submission & Validation

SAS Validation in the Pharmaceutical Industry

Written by and presented by Trish Halley, Software Quality Analyst and Ed Helton, Marketing Strategist.

SAS has historically provided the standard analytic software to the pharmaceutical industry for the analysis and reporting of safety and efficacy data. Title 21, Part II has placed a requirement on the industry to capture the day-to-day compliance of GxPs in a validated system. Validation requires not only a system that is accurate, reliable, and provides the intended performance, but also provides functionality that meets the “predicate rules” of the GxPs.

SAS is continually working with the industry to further assist the drug sponsors in their efforts to migrate, integrate, and validate SAS software. This presentation will review in detail the ongoing efforts at SAS to meet the elements of Part II. It will include details on the following:

- Quality Management System used to produce Part II-enabling software.
- IQ/OQ tools to enable customers to qualify SAS in their environment. This will include a walk-through of one of the self-validating test scripts (written with PROC IML) that is shipped with the tool.
- Migration tools to enable a validated migration to SAS 9.
- Functions and features of SAS to meet the predicate rules.
- MIV project, formed to aid the industry in its migration, integration, and validation efforts.

SAS at PharmaSUG 2004 (continued)

WEDNESDAY, 9:00 AM-9:50 AM ~ Paper Number: SAS2
Room: Diamond II Section: Data Management

Proc Migrate: How to Migrate Your Data and Know You've Done It Right!

Written by David Wiehle and Diane Olson.
Presented by Diane Olson, Senior Systems Developer.

Now that SAS 9.1 is available, what needs to be done to your data libraries in order to make use of the new version? Do you have to migrate your data? If you do, how can you do it most easily? Migrating your data to a new version of SAS presents challenges; depending on the attributes of your data libraries, your migration can be relatively simple or complex. Some issues to consider when adopting SAS 9 include the version of SAS in which your data currently resides, what member types exist in your libraries and whether you must move members from 32-bit libraries to 64-bit libraries. To address these issues, SAS 9.1 includes a new utility procedure, Proc Migrate. Proc Migrate streamlines the process of moving libraries forward to a new release. This presentation introduces Proc Migrate and discusses pitfalls of migrating with the traditional methods - Proc Copy, Proc Cport/Proc Cimport and Proc Catalog. Base SAS also provides tools that can help you guarantee the content and attributes of your data after migration. We discuss how validation tools can be used to validate the integrity of your migrated libraries.

WEDNESDAY, 9:00 AM-9:50 AM ~ Paper Number: SAS3
Room: Topaz Section: FDA Compliance: Electronic Submission & Validation

Defining and Validating CDISC Data Standards to XML in SAS Technology

Written by Anthony Friebel, Edward Helton, Eugenia Bastos, Jane Boone(CDISC) and Michael Kilhuhllen.
Presented by Anthony Friebel, Software Manager, Mike Kilhullen, Technical Architect, and Ed Helton, Marketing Strategist.

We wish to demonstrate the Base SAS technology and tools to use and report investigational data in XML according to the CDISC Operational Data Model (ODM) v1.2 schema and define.XML in Base SAS. The Operational Data Model (ODM) is a vendor neutral, platform independent format for interchange and archive of data collected in clinical trials. The model represents study metadata, data and administrative data associated with a clinical trial. Program fragments used for forward and backward conversions of data from SAS to ODM v1.2 format will be presented, along with displays of the data before and after conversion. The implications of data transmission using this standard will be discussed, including how to display data in ODM v1.2 format, how to validate imported and exported data, and how to otherwise assess conformance of the data with the model. Special attention will be paid to the implementation of the SDM in regulatory submissions to carry out the define.XML process in Base SAS and to demonstrate those processes for the support of CDISC Standards in the most common analysis and reporting software in the industry. Validation of these processes are critical to compliance.

Standard data sets of both efficacy and safety from Phase III trials will be used to demonstrate the conversion of content to format has been performed in compliant manner.

OVERVIEW OF PAPERS

APPLICATIONS DEVELOPMENT

CO-CHAIRS: Lisa McQuay, RTI Health Solutions
David Jemiolo, Dataceutics, Inc.
TIME: Monday, May 24, 8:00 AM - 5:00 PM
Tuesday, May 24, 8:00 AM - 12:00 PM
LOCATION: Diamond I

Paper Number: AD02 *Audience Level*: All
MONDAY, 8:00 AM-8:20 AM Room: Diamond I

A SAS Database Search Engine Gives Everyone the Power to Know

Eugene Yeh ~ PharmaNet Inc.

This application finds the matches to a search parameter that you specify from the contents of all SAS data sets in multiple file directories. The application appears similar to an internet search engine, therefore it is practical for everyone to use.

After you enter a search parameter and click the "GO" button, a summary of the search results is displayed in an HTML or RTF file. The file displays the names of all directories, data sets, variables, and observations that contain a match to the search parameter. By clicking on any data set name, you hyperlink to either the source data set that contains a match or to a newly created data set. This new data set is a subset of the source data set, retaining only observations that contain a match to the search parameter. Advanced features can explicitly include or exclude data sets and variables by name, search for formatted values, and save search parameters.

A SAS database search engine makes "the power to know" available in less time and effort, to more people, thus increasing overall productivity within any organization.

This paper reviews the capabilities of this application, some common uses for it, and the programming techniques implemented.

Paper Number: AD05 *Audience Level*: Intermediate
MONDAY, 8:30 AM-8:50 AM Room: Diamond I

CDX – clinical data management and report programming environment

Gunilla Skollermo ~ AstraZeneca R&D / Pål Edman ~ AstraZeneca R&D

CDX is a SAS application developed by AstraZeneca R&D in Södertälje, Sweden. The CDX (Clinical Data eXplorer) project should provide:

- A flexible development environment facilitating sharing of standards and good practices. The users are SAS programmers who work with data management, creation of clinical study reporting database, study output programming, regulatory defence, and commercial support.
- A facility which creates a standardized reporting database with all derived variables needed for the reports and analysis.
- A production environment which satisfies all regulatory demands regarding traceability and security.

The application should be user-friendly, unobtrusive for SAS-programmers, and appear as an add-on to SAS. The technical setting is a combined network of Windows 2000 and SAS 8.2 on all users' PCs and a server with Windows Terminal Server and Citrix.

The CDX solution includes an interface to Visual SourceSafe and to the Argent Job Scheduler. The user environment is a SAS Tree object overlaid on an ordinary SAS enhanced editor session. Within the tree object CDX specific menus provide all functionality – ranging from Copy and Paste, via Check In and Check Out of versions of program code to Create CDX Database and Declare Clean.

Overview of Papers
APPLICATIONS DEVELOPMENT (continued)

Paper Number: AD15 *Audience Level*: Beginner
MONDAY, 9:00 AM-9:50 AM Room: Diamond I

Building a Powerful Clinical Decision Support System Using SAS/IntrNet.

Nick Paszty ~ XOMA (US) LLC

Simply put, SAS/IntrNet is very clever at providing flexible features to do essentially one thing; web enable your SAS data. This paper will present some helpful key features used to develop a Clinical Decision Support System. Considerations relating to navigation, data presentation, cross study standardization, browser versus file system output destinations and use metrics are presented. In addition to the SAS language, integral technologies like DHTML and JavaScript will be briefly discussed.

Paper Number: AD10 *Audience Level*: Intermediate
MONDAY, 10:00 AM-10:20 AM Room: Diamond I

Tools to Facilitate the Creation of Pooled Clinical Trials Databases

Patricia Majcher ~ J&J Pharmaceutical Research & Development

Data collected from individual clinical trials offer critical information. Of perhaps even greater importance is information that can be gleaned when data from individual trials are pooled. Such pooled databases can be used for several purposes: to answer health authority questions more efficiently, to support publications, and for data mining. However, data from different studies, particularly for long-term drug projects, may utilize different data model structures. Mergers between companies can also introduce variability to a company's data model standards. At Johnson & Johnson Pharmaceutical Research & Development, L.L.C., the Data Warehousing department supports several SAS® macros, which can be used to facilitate the creation and verification of pooled databases. These macros utilize Base SAS® and SAS® Macro Language and can execute on both UNIX and Windows. This paper will present an overview of these macros.

Paper Number: AD18 *Audience Level*: Intermediate
MONDAY, 10:30 AM-10:50 AM Room: Diamond I

Customizing Clinical Data Warehouse for Business Intelligence

Jianming He ~ Solucient / Haibin Shu ~ Barr Research / Sonia Pulgar ~ Solucient LLC

Many companies choose SAS/BASE, SAS/ACCESS as the analytical tool, and data warehousing software other than SAS for data warehouse management. This paper discusses the analyst perspectives in such setting to deal with data loading, extraction and data cleaning. It shows basic methodology in data extraction and data management for achieving the business goals and making strategic business decisions efficiently. SAS users need to work with database administrators effectively to make the data warehouse more user friendly

Overview of Papers
APPLICATIONS DEVELOPMENT (continued)

Paper Number: AD12 *Audience Level*: Intermediate
MONDAY, 11:00 AM-11:50 AM Room: Diamond I

Implementing an Audit Trail within a Clinical Reporting Tool

Gregory Weber ~ DataCeutics, Inc. / Troy Ruth ~ DataCeutics, Inc. / Paul Gilbert ~ DataCeutics, Inc.

This paper is a follow-up to "Overview of a Browser-Based Clinical Report Generation Tool," presented at PharmaSUG 2003. Improved control over the reporting environment was stated as one of the next steps, with the planned addition of a Move to Production facility and accompanying system audit trail. The discussion in this paper will focus on the implementation of the system audit trail.

An audit trail is a useful tool for management to record and review the activities of staff responsible for development and maintenance within a Clinical Reporting Environment. A system audit trail is a vital component in a system intended to satisfy 21 CFR Part 11 and other regulatory requirements.

DataCeutics has developed a system audit trail for the DataCeutics' Report PortalTM (DRP), to further establish the DRP as a comprehensive Clinical Reporting Tool, suitable for use on clinical projects intended for submission. The system audit trail, like the DRP, was developed using the SAS system, SAS/IntrNet application, HTML, and JavaScript.

This paper is an overview of the system audit trail component of the DRP application, focusing on the technical and functional design and compliance with 21 CFR Part 11.

Paper Number: DM10 *Audience Level*: All
MONDAY, 1:30 PM-1:50 PM Room: Diamond I

SASUnit: Automated Testing for SAS

Greg Barnes Nelson ~ ThotWave Technologies, LLC.

Data management is one of the cornerstones of SAS as a language and critical to pharmaceutical research and development. SAS programs that access, manage, analyze and report on data are often taken from vast libraries of tools that are used over and over again for consistency and desirable for their reuse in similar projects. Over time, the number of potential uses of any one program or macro is challenged by the amount of time it takes to test, retest and validate these programs. As these programs become part of the production eco-system in a clinical research environment, it become important that their testability, robustness and manageability become "built-in" to the software development process.

This paper discusses a specific approach to building in that process to each and every program to monitor the conditions SAS programs encounter and proactively test for and announce any validation issues. We will explore the concept of automated tests through assertions, events and their attributes, event status management, and automatic notification of events to interested parties. These concepts are presented from the perspective of the SAS programmer and the systems analyst.

Overview of Papers
APPLICATIONS DEVELOPMENT (continued)

Paper Number: AD08 *Audience Level*: Advanced
MONDAY, 2:00 PM-2:20 PM Room: Diamond I

COMPAREXLS - a powerful comparison tool for spreadsheet difference resolution

Kirill Tchernakov ~ Boehringer Ingelheim

Boehringer-Ingelheim (BI) as other pharmaceutical companies has the need to report the safety profile of a drug to regulatory agencies. In an effort to address this need as quickly as possible and in a manner that allows the pooling of data from all phases of the drug development process, BI has developed the "Extensible Project Database Loading and Table Translation Program" (XPDL) process. XPDL uses SAS to remap and combine multiple data sources into one dataset. The process is dynamic and table driven. The tables are edited in Excel® and are maintained by Data Managers.

The recent enhancement to XPDL was the development of the CompareXLS SAS macro program. This macro was written to ease the burden of comparing the standard table layouts (spreadsheets) and resolves the difference between them. The spreadsheets are metadata that describes the source-to-target relationships of the project database and serves as an input for building project databases by the XPDL 'load program.'

The CompareXLS converts the spreadsheets into datasets and compares them by their variables and attributes, offering different ways of problem resolution. After the differences are resolved, CompareXLS can effectively transform the resulting dataset into a new spreadsheet. The saving of time and quality are tremendous.

Paper Number: AD09 *Audience Level*: Beginner
MONDAY, 2:30 PM-2:50 PM Room: Diamond I

CheckAll: a routine to produce a frequency of response data set from multiple-response data

Ronald Fehd ~ Centers for Disease Control

Multiple-response data from survey questionnaires where questions have the instruction "check all that apply" present a challenge to the SAS software programmer. For a simple question, the answer may be either A or B; the sum of percent response is 100%. For the series of variables in a multiple-response question, the answer may be both A and B; because response rate for each variable in the series is dependent upon the other variables, the sum of the percent response may be greater than 100%.

This paper examines the SAS software proc FREQ output data set and discusses the construction of a standardized data set containing frequency of response information for multiple-response data for graphics presentation.

Overview of Papers
APPLICATIONS DEVELOPMENT (continued)

Paper Number: AD19 *Audience Level*: Beginner
MONDAY, 3:00 PM-3:50 PM Room: Diamond I

Using SAS Drug Development as a Report Management Application

Barry Cohen ~ Planning Data Systems, Inc.

Many statisticians and statistical programmers in the pharmaceutical industry will first come to know SAS Drug Development as a product that addresses their regulatory-compliance issues (auditing, versioning, and security) as they develop their on-going analysis programs, data, and documents for NDA filings. However, the product provides a full, flexible processing environment that can be used in other ways. In this paper, I examine standard features of SAS Drug Development that allow it to serve as a Web-enabled report management application for a library of SAS-based report programs. The application covers typical functions such as (1) loading SAS macro report programs to the environment; (2) building report parameter-solicitation screens; (3) providing a user interface to select report programs, set report parameter values, and execute reports; (4) managing report output sets and providing a user interface for output file viewing. Emphasis is on what SAS Drug Development can provide for a report management application out of the box without a user-customization of the application through the product's API.

Paper Number: AD11 *Audience Level*: All
MONDAY, 4:00 PM-4:20 PM Room: Diamond I

Organizing and Building a Centralized SAS Macro Library

Snow Fu ~ Merck & Co. / James Wu ~ Merck & Co.

In many pharmaceutical companies, writing and sharing SAS macros has become common practice in the statistical programming area. To properly manage this proliferation of SAS macros, it seems more and more necessary for a pharmaceutical company to build and maintain its own centrally controlled SAS macro library. These libraries may contain many types of macros, including statistical, utility, text-handling and data-manipulation macros. In this paper, we will discuss some suggestions for effectively organizing and building a centralized SAS macro library.

Paper Number: AD01 *Audience Level*: Intermediate
MONDAY, 4:30 PM-4:50 PM Room: Diamond I

A Macro Generating a Report-Ready Table in Different Layout and Format

Renda Cheng ~ Bristol-Myers Squibb Company

Generating comprehensive and report-ready tables for clinical studies has been very challenging to SAS programmers. In order to meet this need a macro called Table Macro was developed in the Oncology Group of BMS two yeas ago. Since the macro went into production, the quality of the tables has been highly recognized by our customers (clinicians and statisticians). It has saved a lot of our programming time and shortened the process of application for NDA.

The macro has the following features:

(1) EASY TO USE. Users only need to enter a few parameters to the macro to generate a table. These parameters are very straightforward and easy to understand and remember.
(2) VERY FLEXIBLE AND POWERFUL TO GENERATE OUTPUTS IN A DESIRABLE LAYOUT AND FORMAT. Users can easily generate a frequency table or a statistic table or both in one table. Users can also easily add categories in the table and control denominators, number of columns/column groups, decimal, column width / alignment / label... etc. Users even can add their own code to the macro to customize the table layout and format in order to meet clinical study needs.
(3) EASY TO DEBUG. The macro generates error messages to guide the user when incorrect specifications are encountered.

Overview of Papers
APPLICATIONS DEVELOPMENT (continued)

Paper Number: AD13 *Audience Level*: All
TUESDAY, 8:00 AM-8:20 AM Room: Diamond I

Version Control on the Cheap. A User-Friendly, Cost-Effective Revision Control System for SAS

Tim Williams ~ PRA International

Revision control, an integral component of code development in many languages, is often absent from SAS programming. Cost and ease of use are often cited as two of the primary reasons why a revision control system is not in use. However, the need for such a system is growing. Producing an audit trail for changes to code is becoming increasingly important within the pharmaceutical industry. From a programmer's perspective there are many advantages to a system that can easily differentiate between programs that are under development, have been validated, or were used for specific production runs (such as an interim analysis or final analysis). The ability to roll back changes during development and easily compare revisions to files can greatly increase programming efficiency.

This paper describes the implementation of free, open source components that comprise an easy to use revision control system for SAS programs. The system is deployed in a large organization where SAS is used for the analysis of clinical trials data. The following discussion is intended for programmers at any experience level who want to better manage their SAS programming projects.

Paper Number: AD07 *Audience Level*: All
TUESDAY, 8:30 AM-8:50 AM Room: Diamond I

Mergers, Consolidations and Proc Compare

Fang Dong ~ Pfizer

This paper discusses how to verify data integrity after database consolidations. System and database consolidations usually follow corporate mergers and acquisitions. A critical need in the database consolidation process is to ensure the integrity of the data. In the heavily regulated pharmaceutical industries, it is also important to ensure reproducibility of analysis and report output. This paper will describe how to use SAS proc compare and macro utilities to help meet these needs. Detailed explanation and examples will be provided to demonstrate how to construct a SAS program to check large amounts of data.

Paper Number: AD04 *Audience Level*: All
TUESDAY, 9:00 AM-9:50 AM Room: Diamond I

Web Services Integration and the SAS Drug Development Platform

Chris Olinger ~ d-Wise Technologies, Inc.

The SAS Drug Development platform is an extensible, web-based application for the hosting, processing, and management of clinical and pre-clinical data. The system is 21CFR Part 11 compliant, runs over the Internet, integrates with SAS, and provides direct interfaces for controlling the various components of the system. The system employs Java EJB technology, Web Services, and a WEBDAV interface to deliver Internet-based clinical functionality, and to provide integration points for controlling basic functionality such as auditing, uploading, downloading, and versioning. This presentation covers the ideas surrounding Web Services and reviews the main methods for integrating with SAS Drug Development (as there are several). The presentation is intended for people with Java/Web experience, and for SAS programmers that would like to see what you can do with the system.

Overview of Papers
APPLICATIONS DEVELOPMENT (continued)

Paper Number: AD16 *Audience Level*: Intermediate
TUESDAY, 10:00 AM-10:20 AM Room: Diamond I

A Quicker Way to Inputting and Sorting Large Dataset Using SAS

Lin Yan ~ MedFocus

This paper is concerned about the speed of inputting and sorting large datasets using SAS. Usually people use the only one SAS procedure Proc SORT to input data and sort data as well. However, this procedure is not efficient when the dataset is large. A two-step method by which first the data is input using DATA step and then is sorted using procedure SORT can enhance the speed of inputting and sorting a large dataset. Experiments on large clinical datasets show that the two-step method reduces the run time greatly

Paper Number: AD17 *Audience Level*: Intermediate
TUESDAY, 10:30 AM-10:50 AM Room: Diamond I

The Big Button Approach for Quickly Running and Printing Large Numbers of Tables

Mary Cowmeadow ~ BioStat Reports

To meet speed goals, it is important to produce all study tables as fast as possible at certain points during the study and at final database release. Traditionally one table had been run and printed at a time, a slow and tedious manual process. A much quicker method was needed. Also the work team needed to slice and dice, pool and collapse, integrate and subset in order to look at the data in many ways for an exploratory study.

The “Big Button” driver was developed to help meet these goals. With one “submit”, all programs would run. In SAS the Output Delivery System and Proc Tabulate were chosen because of their slice and dice capability and nice output. As a consequence, the study workgroup, rather than asking for more and more, actually decided instead to reduce the number of tables quite substantially. Finally, Microsoft Binder, a part of Microsoft Office which few people seem to know they have, was a useful way to print large numbers of tables. One can drag and drop many tables into a binder in one step and then e-mail it to a central center for rapid printing of multiple copies.

Overview of Papers

CODER'S CORNER

CO-CHAIRS: Cindy Song, Merck & Co., Inc.
Jim Edgington, Quintiles, Inc.
TIME: Monday, May 24, 3:00 PM - 4:45 PM
Tuesday, May 25, 8:00 AM - 11:00 AM
LOCATION: Topaz

Paper Number: CC09 *Audience Level*: Intermediate
MONDAY, 3:00 PM-3:10 PM Room: Topaz

Changing that pesky datetime to a date

Stanley Fogleman ~ HARVARD CLINICAL RESEARCH INSTITUTE

Databases and spreadsheets commonly store both the date and the time in a single field. Often, the time field is initialized to midnight. Here is a simple macro that can change a datetime to a date.

Paper Number: CC12 *Audience Level*: Intermediate
MONDAY, 3:15 PM-3:25 PM Room: Topaz

Use Regression Algorithm Identifying Repeated Use Patients

Haibin Shu ~ Barr Research

For repeated use study, a patient can enroll in multiple times with each time assigned different patient number. In many situations, a patient is asked to provide any used patient numbers but give only partial answers, i.e., several most recently used numbers. In order to do analysis on a per patient base, it's important to link all those enrollment numbers associated with a patient. In a study with large population, the way a patient provides old enrollment patient numbers can be quite different from one to another and the task to link them can be very challenging. In this paper, the author designed a systematic regression algorithm to achieve the task and provided codes in SAS to implement it.

Paper Number: CC01 *Audience Level*: Intermediate
MONDAY, 3:30 PM-3:40 PM Room: Topaz

A technique to create a Numeric Variable from a Character Variable

Hany Aboutaleb ~ BiogenIdec

If you have a character variable in your dataset, and you wish to transfer this variable to numeric variable instead, this article can help you. I will describes and illustrate by example how to use Proc Sort, and Format with the help of SAS Macro facility to create a new numeric variable in your data set from your character variable.

Paper Number: CC07 *Audience Level*: Beginner
MONDAY, 3:45 PM-3:55 PM Room: Topaz

Using Compute Blocks in Proc Report

Sharon Dunn ~ Sepracor

Proc Report is widely used to generate tables and listings. By incorporating the use of compute blocks within the Proc Report, we can utilize data null capability to get around some common challenges. For anyone still using a version of SAS 8.1 or earlier, there is a 10 line text limitation for titles and footnotes. Compute blocks will allow for as many lines of text as needed, as well as letting you create page specific titles and footnotes. They can also be a very useful way of specifying varying category titles within the report body, particularly if the data dictates which titles should be used. Another advantage of using compute blocks is to create the line 'Number of Patients Dosed' without having to add an observation to your dataset. You simply calculate macro variables for these N's and use a compute block to output the information into the report.

Overview of Papers
CODER'S CORNER (continued)

Paper Number: CC16 *Audience Level*: Intermediate
MONDAY, 4:00 PM-4:10 PM Room: Topaz

Automating the Documentation of an Oracle Clinical Database Specification

Kyle McBride ~ Instat Consulting, Inc.

This paper will demonstrate how custom study documentation can easily be generated by using SAS to tap into a clinical trial database system such as Oracle Clinical. With an understanding of the Oracle Clinical database structure, what sometimes is a manually compiled (and labor intensive) document can be automated using SAS/Access to Oracle, Proc SQL and ODS. An example is given for a database structure document. Advantages over other reporting options or existing tools are mentioned.

Paper Number: CC18 *Audience Level*: Intermediate
MONDAY, 4:15 PM-4:25 PM Room: Topaz

Using a Decision Table to Program Complex Date Logic

Arthur Collins ~ Biogen Idec, Inc.

Programming tasks involving complex date logic, such as to determine treatment emergence of adverse events or medication concomitance, can be time consuming and prone to error. This paper will propose and demonstrate an approach utilizing a decision table, which systematically handles all logical possibilities. This approach promises to make the programming easier and more efficient, while offering more control to the study statistician. The statistician can be the one to set up the table, allowing them to ensure accuracy and correctness, as well as being able to easily make changes to the logic when necessary. Finally, the statistician can use the table to define illogical values, such as a resolution date prior to onset, which they would like brought to their attention for review.

Paper Number: CC20 *Audience Level*: All
MONDAY, 4:30 PM-4:40 PM Room: Topaz

PROC POWER: V9.1

Deborah Bauer ~ Sanofi-Synthelabo

The features that were added to SAS Proc Power in V9 will allow many users to do all their power and sample size calculations in SAS. This free SAS proc will allow companies to save money by canceling their licenses for software packages that only do sample size calculations.

Not only is this upgraded proc very full featured, but the ODS capabilities in SAS make it easy to take the results of the calculation directly into a finished document.

This paper intends to compare packages, at such a detailed level, that readers, can use this article as the structure for a project to justify the discontinuation of other software licenses.

Overview of Papers
CODER'S CORNER (continued)

Paper Number: CC17 *Audience Level*: Intermediate
TUESDAY, 8:00 AM-8:10 AM Room: Topaz

Module to create global Macro Variables with distinctive names matching Date values they carry

Anatoly Kulinsky ~ FBF, Fleet Credit Card Services

Macro Function is designed to generate various uniquely named Global variables to hold date values corresponding to the names. Macro variables created automatically for future use in SAS code. Available arguments include date formats and time intervals, alignment within the intervals, explicit starting point or default system date, number and direction of iterations –up to seven parameters with assigned default values. This flexibility allows using same tool with different arguments to produce multiple macro variables of different date format representing essentially same date, if requires.

The INTNX function is used to shift time interval. CALL SYMPUT routine uses date values to create macro variables. Implementation allows referencing and comparing dates stored in the variables or embedded into various SAS names, external file or database names. It could be employed to create time-series suffixes or custom formats allowing automation and summarization of data by applying combination of two different date formats on the same data.

Function is designed as an application for users. Error Handling introduced to exit program using %GOTO and %LABEL statements with explanatory ERROR messages to the SAS LOG achieving self-tutorial effect. It is equipped for job control and produces Return Codes permitting to identify even multiple exceptions.

Paper Number: CC19 *Audience Level*: Intermediate
TUESDAY, 8:15 AM-8:25 AM Room: Topaz

GSUBMIT: Simple Customization of your SAS Application Toolbar in SAS for Windows

Robert Howard ~ Synteract, Inc.

Have you ever wanted to delete all your working SAS datasets without closing your SAS session or without having to type "proc datasets kill; run; quit;"? Do you have a SAS program that you run often but you don't like to search for it in your directories? Ever wished you could execute a proc print without typing? Well, there is a simple way to add a customized button to your Application Toolbar that will run a specific SAS program or SAS code – all at the click of your mouse.

Paper Number: CC04 *Audience Level*: All
TUESDAY, 8:30 AM-8:40 AM Room: Topaz

BACKUPS AT THE CLICK OF THE MOUSE

Jim Box ~ RTI International / Joey Morris ~ RTI International

Creating a backup of SAS programs, output files and datasets is an important task for any computing project. Unfortunately, it is also one of the more tedious and error-prone tasks. This paper will show you how to configure your PC to automatically perform these backups (create indexed archives) with one click of the mouse.

Overview of Papers
CODER'S CORNER (continued)

Paper Number: CC05 *Audience Level*: All
TUESDAY, 8:45 AM-8:55 AM Room: Topaz

Integration of Microsoft Office and SAS Software: Implementation of GUI and Automation

Rubin Nan ~ PRA International / David Mullins ~ PRA international

This paper describes techniques that use MS-Office embedded Visual Basic tools or the Visual Basic software to create a Graphical User Interface (GUI). Using this interface, users can direct the SAS© session, DATA steps, and SAS procedures to run. They can also input SAS macro parameters (macro variable values) from a front end to resolve macro variables in SAS programs.

To clarify our description, we use simple examples to demonstrate how to create a GUI, to set up component properties, and to write the Visual Basic and the SAS code to implement the integration and automation.

Paper Number: CC06 *Audience Level*: Intermediate
TUESDAY, 9:00 AM-9:10 AM Room: Topaz

%ArrayPerm: A SAS Macro for Permutation Analysis of Microarray Data

Deqing Pei ~ St. Jude children's research hospital / Wei Liu ~ St.Jude Children's research hosptial / Cheng Cheng ~ St.Jude Children's research hospital

Microarray has become a common tool for identifying genes of interest that are differentially expressed under different biological conditions. Software packages such as Spotfire, GeneSpring, and Xcluster are commonly used in research laboratories. These packages provide a set of analysis tools for accessing, visualizing, normalizing and filtering data with modest learning curve for biologists without little statistical background. However, wide applications of DNA microarray technology in gene discovery, disease diagnosis, pharmacogenomics etc. require more sophisticated statistical analysis methods beyond simply looking for genes up- or down-regulated, thereby requiring solid data modeling on the basis of valid statistical methods. When the expression data or their transformations cannot be assumed to distribute normally, or the investigator wants to assess the reliability of gene-selection procedures, permutation tests are often desired. Software packages mentioned above do not have flexibility or capability for implementing permutation tests under sophisticated models such as general linear models, generalized linear models, GEE, Cox regression models, etc. Here, we present SAS macros to facilitate permutation analysis on a broad spectrum of statistical models that can be chosen according to specific objectives of a microarray experiment. This macro is a very flexible tool for one to implement permutation tests quickly and easily.

Paper Number: CC08 *Audience Level*: Intermediate
TUESDAY, 9:15 AM-9:25 AM Room: Topaz

Create all those tables and call PROC SQL once!

Stanley Fogleman ~ Harvard Clinical Research Institute

Many beginning programmers don't realize the overhead in calling PROC SQL repeatedly - by separating the first, inbetween and last calls to PROC SQL, a great deal of overhead can be saved - result: job runs in much less time.

Overview of Papers
CODER'S CORNER (continued)

Paper Number: CC10 *Audience Level*: All
TUESDAY, 9:30 AM-9:40 AM Room: Topaz

Cutting the SAS LOG Down to Size

Malachy Foley ~ University of North Carolina

Looking through a large SAS LOG (say 250 pages)for NOTE's and WARNING's that might indicate a problem with your SAS program is necessary, but no fun. This paper presents a SAS program which distills the SAS LOG to list only the critical messages in the LOG. The use of the program increases program reliability and programmer productivity.

Paper Number: CC13 *Audience Level*: All
TUESDAY, 9:45 AM-9:55 AM Room: Topaz

Frequency Generation Using Indicator Variable

Jinshi Zhou ~ Quintiles

A quick and easy way of generating entries for frequency table is presented. It involves the construction of indicator variables that correspond to those properties in the table in data steps and a call of PROC UNIVARIATE from SAS on these indicator variables to generate a data set that contains all of the key quantities in the table. It then uses these key quantities in a data step to create all entries for the table. Finally the data set is presented in the way requested via PROC REPORT or DATA _NULL_.

Paper Number: CC14 *Audience Level*: Intermediate
TUESDAY, 10:00 AM-10:10 AM Room: Topaz

Creating a Mixed-case Format for Mono-case Data Using an Input Control Data Set with Proc Format

Lisa Soberano ~ PPD Development, Inc.

PROC FORMAT is a useful Base SAS tool in displaying data in a more preferable way, without altering the data itself. Applying formats to our data is like using "make-up". We want the data to look its best, whether presenting data to a client or internally.

This article assumes a basic familiarity with the DATA step and PROC FORMAT. We will review a commonly overlooked function of PROC FORMAT, which is how to create a format using an input data set (the CNTLIN= option on the PROC FORMAT statement), as well as, how to output formats to a data set (CNTLOUT= option). The example application in this article is also one that will come in handy for displaying large amounts of mono-case characters. We will create a mixed-case format for character data that is stored in either all upper or all lower case. Although we have the use of the UPCASE and LOWCASE functions available through SAS, the code covered in this text will give you a starting point to developing your own mixed case function.

The SAS functions that are incorporated in the mixed case example are UPCASE, LOWCASE, TRANWRD, SCAN, LENGTH, SUBSTR, TRIM, and COMPRESS.

Overview of Papers
CODER'S CORNER (continued)

Paper Number: CC15 *Audience Level*: Beginner
TUESDAY, 10:15 AM-10:25 AM Room: Topaz

Common Error: Wrong result

Varsha Shah ~ Univ of North Carolina

The routine, frequent use of functions, internal automatic variables, etc. creates such familiarity that it becomes second nature and forges a friendship. This in turn creates such confidence that once in a while there is a relaxed approach toward checking the results. It is extremely important to emphasize that the above mentioned tools in SAS are very powerful, but what could taint the results is the concurrent environment. This paper illustrates two examples of how your algorithms could lead to erroneous conclusions - 1) when the LAG function is used improperly, and 2) when you have irregular data.

Paper Number: CC02 *Audience Level*: All
TUESDAY, 10:30 AM-10:40 AM Room: Topaz

Using SAS to Speed up Annotating Case Report Forms in PDF Format

Dirk Spruck ~ Covidence GmbH / Monika Kawohl ~ Covidence GmbH

One of the first things programmers do before they start working on clinical data is to create an annotated case report form (CRF). Variable names are written on the CRF to provide a link between the fields on the form and the variables in the data sets. Annotating the CRF electronically increases the readability and reusability of the document. For electronic submissions the Food and Drug Administration (FDA) requires a completely annotated CRF in PDF format.

This paper describes the technology used to place annotations in a PDF file and the process of copying annotations between pages. SAS is used to write the variable names from specified data sets and the associated formats as notes in Forms Data Format (FDF). After the CRF is converted to a PDF file the notes are imported into the PDF document and manually placed in the appropriate position on the CRF page. For repeat pages the notes are exported. Using SAS the exported notes are copied and the page numbers are replaced. The notes are then re-imported into the PDF document.

Paper Number: CC03 *Audience Level*: All
TUESDAY, 10:45 AM-10:55 AM Room: Topaz

A Simplified Way to Create Flags for the Three-Dimensional Prior, Concomitant, and Post Medication

Aileen Yam ~ Aventis, Inc.

A three-dimensional classification of non-study medication into prior, concomitant and post medication categories was proposed at CDISC. Please see Section I of this paper for the classification. This paper presents a simplified way to create flags for each condition. This paper also puts forth for consideration a conservative method of imputing partial dates.

Overview of Papers

DATA MANAGEMENT

CO-CHAIRS: David Izard, GlaxoSmithKline
Cecilia Mauldin, PPD Development, Inc.
TIME: Tuesday, May 25, 3:30 PM - 5:30 PM
Wednesday, May 26, 8:00 AM - 12:00 PM
LOCATION: Diamond II

Paper Number: DM04 *Audience Level*: Intermediate
TUESDAY, 3:30 PM-3:50 PM Room: Diamond II

Helpful Hints on Developing a User Friendly Database with SAS/AF

Sy Truong ~ Meta-Xceed, Inc.

Developing an effective database application requires an interface that is easy for the user. This paper will explore the features of SAS/AF in SAS version 8.2 and methodologies of building a successful database. It combines user interface suggestions for the front end while also suggesting back end SCL, SQL and data step logic that makes the software efficient to program and to operate. The majority of the examples are technical tips but there are also shared lessons learned from collaborating with end users which prove to be very important in creating an effective application.

Paper Number: DM09 *Audience Level*: Intermediate
TUESDAY, 4:00 PM-4:20 PM Room: Diamond II

Use That SAP to Write Your Code

Sandra Minjoe ~ Genentech, Inc.

In this electronic age we live in, we usually receive the detailed specifications from our biostatistician in an electronic statistical analysis plan (SAP). We often then turn around and retype much of these specs into our programs.

This paper describes and gives examples of how we can take those electronic specs, from, say, an Microsoft Excel spreadsheet or Microsoft Word table, and convert them into a dataset or code logic. This prevents us from introducing typos because we no longer have to retype information already provided. The real savings come later on, though, when the specs change (as they so often do). By using the techniques outlined in this paper, when specs change we simply need to rerun, rather than recode.

Paper Number: DM05 *Audience Level*: All
TUESDAY, 4:30 PM-4:50 PM Room: Diamond II

The use of Automatic ODMs on late phase trials to increase data quality while decreasing costs.

Kathleen Kushner ~ Covance Periapproval / Philip Pellicone ~ Covance Periapproval

Many late phase and Post-Marketing clinical trials utilize Obvious Data Modifications (ODMs) to correct hard to clean data. ODMs are database updates that are self-evident and pre-approved by sponsors/investigators, and can therefore be made without querying a site. In many late-phase trials, patient Quality of Life and diaries are completed. Since the patient captures this data, it is usually very dirty and not able to be queried. The utilization of approved ODMs allows this data to be corrected resulting in an effective statistical analysis.

The manual process of updating large amounts of data is very time consuming and has a high risk for human error. The time, and ultimately costs associated with Data Management staff manually updating potentially thousands of ODMs does not seem a feasible process to continue. Automatic ODMs would serve by creating better quality data in less time. This presentation will discuss the process of automatically updating ODMs utilizing SAS as an update, validation and audit trail tool. Data Management considerations and a cost savings analysis will also be introduced.

Overview of Papers
DATA MANAGEMENT (continued)

Paper Number: DM03 *Audience Level*: All
TUESDAY, 5:00 PM-5:20 PM Room: Diamond II

"Haven't We Met Somewhere Before?" and Other Useful Lines

David Rucker ~ BioCor

"Know Thy Data". As programmers, we see all types of data, in all forms, on all media, from different operating systems, from different customers, and possibly in different languages. We may never agree on the approach used in our code. We may never agree on programming standards. We may never agree on data presentation. Nonetheless, any user of the SAS System will tell you this is what he appreciates most about the product. Surely, the one thing we can all agree on is that you MUST know your data. The first step in this new relationship is to introduce yourself.

In this paper, I will discuss several fundamental methods and possibly some new ideas and concepts to facilitate learning about the data. Although the intended audience for this paper is the novice programmer, it may provide the more seasoned programmer with new ideas and tools. (Base SAS, Microsoft Windows operating system).

Paper Number: DM06 *Audience Level*: Advanced
WEDNESDAY, 8:30 AM-8:50 AM Room: Diamond II

Invalid: a Data Review Macro Using PROC Format Option Other=Invalid to Identify and List Outliers

Ronald Fehd ~ Centers for Disease Control

Data cleansing, or as it is more euphemistically known: data review, often occupies too much of a programmer's time and energy. With a properly written data dictionary, a data set will contain appropriate formats for each variable; one can then cut and paste the format definitions into exttt{PROC Format} value statements and label all outliers with the value statement option: exttt{other = "Invalid"}. This routine combines a format catalogue and the contents of a data set, then uses that information to write data steps which select all outliers. Reports are written, by identifier, for print review and to file, for later use as an exttt{\%include} for updating purposes. Expected audience is intermediate and advanced programmers and macro users.

Paper Number: DM07 *Audience Level*: All
WEDNESDAY, 10:00 AM-10:20 AM Room: Diamond II

Why SAS is the Best Place for Your Clinical Data

Steven Wilson ~ MAJARO InfoSystems, Inc.

The SAS System contains an incredible amount of functionality for managing your data files as a database. In most respects, including data storage, data access, and data management, SAS can be considered a database. This paper will review the database functionality available within the SAS System and how this relates to the task of clinical trails data management. Sample code, which illustrates many of these database techniques, is provided.

Overview of Papers
DATA MANAGEMENT (continued)

Paper Number: DM02 *Audience Level*: Intermediate
WEDNESDAY, 10:30 AM-10:50 AM Room: Diamond II

Constructing Analysis Files: A Multi-Tiered Approach

Tyler Cole ~ Pacific Data Designs, Inc.

Data for clinical studies are often organized into analysis files which are typically used for statistical analyses as well as for generation of tables, figures, and listings. Analysis files generally have a number of user-created and non-user created variables shared in common. As a result, programs used to create these files may contain copies of the same code. If many analysis files are needed, uncontrolled code duplication becomes a significant problem impairing the consistency and maintenance of programs. A system design able to control redundant code is thus desired.

This paper, intended for intermediate to advanced users of Base SAS, describes one innovative programming approach which can be implemented to organize and contain the data steps and SAS procedures used to create analysis file variables. With this approach, SAS statements are grouped into function-specific macros stored separate from analysis file programs. These macros are further grouped into tiers each representing a support level – higher-tiered macros inherit data steps and SAS procedures contained in lower-tiered macros. Analysis file variables are created through referencing and executing tiered macros; as a result, duplicate code is controlled, programs are easier to maintain, and data integrity is enhanced.

Paper Number: DM01 *Audience Level*: Intermediate
WEDNESDAY, 11:00 AM-11:20 AM Room: Diamond II

Data Summary at a Glance

Varsha Shah ~ Univ. of North Carolina

This paper introduces techniques for producing documentation, in the form of summary tables, to accompany public use data sets. These tables must be comprehensive enough to serve their purpose of enabling potential data users to quickly get a feel for the data, yet the production technique must enable processing of many data sets in an automated fashion. This automation is accomplished via use of macro features. A highlight of the presentation is a technique for displaying multiple frequency tables on a page in a columnar layout.

Paper Number: DM08 *Audience Level*: Advanced
WEDNESDAY, 11:30 AM-11:50 AM Room: Diamond II

A SAS Macro for Rapid Identification of Missing Data Records

Kim Truett ~ independent

“Does Your Database Contain all Expected Records for each Subject?” The comprehensive review of clinical trial databases should include verification that each dataset contains all expected records for each subject/visit. Because this verification step is time consuming and missing data can be difficult to detect, it is often omitted or done inconsistently. However, this verification step is critical to ensure that the database is complete prior to locking the database so that that statistical analysis can begin.

To facilitate this review, I developed a generic macro which outputs a graphical representation of the data records present for each subject at each visit. The macro produces "one grid per visit" indicating the presence or absence of a record within every dataset for each subject. This graphical format simplifies review and allows for rapid identification of all missing records.

Overview of Papers

FDA COMPLIANCE: ELECTRONIC SUBMISSION & VALIDATION

CO-CHAIRS: Shahn Hawkins, Quintiles, Inc.
Jim Baker, Amgen, Inc.
TIME: Tuesday, May 25, 11:00 AM - 5:30 PM
Wednesday, May 26, 8:00 AM - 12:00 AM
LOCATION: Topaz

Paper Number: FC03 *Audience Level*: Intermediate
TUESDAY, 11:00 AM-11:20 AM Room: Topaz

Automation of Patient Narratives

David Izard ~ GlaxoSmithKline / Eric Simms ~ GlaxoSmithKline

FDA and ICH guidance documents call for the submission of subjects' study experience in narrative form for those subjects who meet specific criteria. This narrative text is typically developed manually utilizing word processing technology and information sourcing from patient line listings or individual case report form tabulations. Historically this is a manual, labor intensive process requiring significant human resource and lengthy timelines.

This paper will explore the use of SAS to identify patients requiring narratives, and automate the development of patient narrative text and make it available to clinical colleagues in a readily usable form, thus reducing the effort and time required to produce this regulatory requirement.

Paper Number: FC01 *Audience Level*: Intermediate
TUESDAY, 11:30 AM-11:50 AM Room: Topaz

XML Data Mapping under SAS 9 - From the Perspective of Electronic Submissions

Tianshu Li ~ Merck Co. & Inc.

Like most other government agencies, the Food and Drug Administration (FDA) is adopting extensible markup language (XML) as one of the standard data formats for electronic submissions. Nevertheless, since SAS will still remain as the major analytical tool in the foreseeable future, the integration and compatibility between XML and SAS software are becoming critical in adopting XML for electronic submissions. The switch from traditional SAS Xport data format to XML presents opportunities because of the huge potential of XML for web- based technologies, and faces new challenges because of the fundamental differences in data structures between them. This paper focuses on the testing and evaluation of the mapping process between SAS data and XML. In addition, the readiness of the SAS 9 XML engine for electronic submissions is also assessed. Future improvements and suggestions are also discussed.

Paper Number: FC11 *Audience Level*: Advanced
TUESDAY, 1:30 PM-1:50 PM Room: Topaz

Critical Success Factors in Pharmaceutical Innovation

Thomas Burger ~ Eli Lilly and Company / Michael Lajiness ~ Eli Lilly and Company

Technology is one of the most important factors that will influence which pharmaceutical companies prevail in the information age. Effective, efficient and timely implementation of tools that speed information delivery is a primary driver of competitive advantage. However, difficulty often arises in determining how to adopt tools and integrate them into the business. Interruptions to the business and failures in software acquisition are frequently costly, in terms of money, resources, and lost opportunities necessitating an innovation-based approach. We discuss strategy and considerations essential to matching SAS technologies with key business drivers at the corporate scale. In addition, we will discuss our vision for future software opportunities that would significantly enhance pharmaceutical research.

Overview of Papers
FDA COMPLIANCE: ELECTRONIC SUBMISSION & VALIDATION (continued)

Paper Number: FC05 *Audience Level*: All
TUESDAY, 3:00 PM-3:20 PM Room: Topaz

Build Portable Structures and Programs for Electronic Regulatory Submissions

Wei Cheng ~ Isis Pharmaceuticals

The Food and Drug Administration (FDA) continues to update a series of guidance documents on electronic regulatory submission as the technology evolves. The FDA provides recommendations and examples for the procedures, structures, and file formats for sending electronic submissions. Most of the companies in the pharmaceutical industry are moving toward their goals of submissions in electronic format in place of paper. This paper will discuss how to build portable structures and programs to support efficient regulatory submission in electronic format. It is intended for SAS programmers of all skill levels, and for SAS Systems on different operating systems.

Paper Number: FC06 *Audience Level*: Theoretical
TUESDAY, 3:30 PM-4:20 PM Room: Topaz

Using Cyclomatic Complexity to Assess Test Coverage for SAS Programs

Michael Harris ~ Amgen Inc.

The Cyclomatic Complexity metric described by Watson and McCabe provides an objective measure of the complexity of a given module of program code by examining its decision structure. This article examines issues related to using the metric. These include identifying relevant SAS flow control constructs, creating flow graphs of SAS programs, and methods for deriving test cases from basis path analysis. Emphasis is placed on programs typical of the biotechnology and pharmaceutical industry, but the principles are applicable to any kind of program.

Paper Number: FC07 *Audience Level*: All
TUESDAY, 4:30 PM-4:50 PM Room: Topaz

A Useful Macro for Converting SAS Data sets into SAS Transport Files

Xingshu Zhu ~ Merck company / Shuping Zhang ~ Merck Company

In 1999, the FDA issued a guidance imposing certain requirements on electronic submissions. Under this guidance, an analysis data set must be converted into a SAS transport file. As a result, it is important for programmers to have a convenient and reliable tool to perform this conversion. We have developed such a tool in the form of a utility macro. This macro can also be used to efficiently and effectively split data sets when they exceed the FDA-imposed size limitation. We have tested this macro in many projects, and it has proved to be a very efficient and advantageous tool in FDA submissions.

Overview of Papers
FDA COMPLIANCE: ELECTRONIC SUBMISSION & VALIDATION (continued)

Paper Number: FC10 *Audience Level*: All
TUESDAY, 5:00 PM-5:20 PM Room: Topaz

eCTD -- a New Standard for FDA Electronic Submission.

Shawn Wang ~ MedXview Inc

The future is here, eCTD is playing a more and more important role in FDA electronic submission. The eCTD can be seen as an Open Standard. It is vendor-neutral for submission assembly. Any eCTD-compliant submission prepared in any compliant environment can be transferred to another environment built on technology from another vendor. FDA reviewer can now easily use EVS processor performs rigid validation of backbone against DTD which requires strict adherence to specifications and do not add or modify leafs within the backbone. Once a submission is sent in eCTD format all future submissions for the application should be in eCTD format:

What doesn't change?
Data files submitted in SAS XPORT format.
Documents submitted in PDF Format.
Acrobat 5 is available to all reviewers.
Draft labeling submitted in MS Word.
Office XP is current standard.

After careful demonstration of eCTD in this paper, pharmaceutical companies can find the underlying rules and patterns and thus help both management and regulatory authority make better decisions in the FDA electronic submission. This paper presents the challenges and solutions for successful submission of eCTD.

The paper is intended for all level of SAS users. Currently the system runs on PC SAS on Window XP.

Overview of Papers
FDA COMPLIANCE: ELECTRONIC SUBMISSION & VALIDATION (continued)

Paper Number: FC02 *Audience Level*: Intermediate
WEDNESDAY, 8:00 AM-8:50 AM Room: Topaz

Data Standards with and without CDISC

Sy Truong ~ Meta-Xceed, Inc.

Data standards can make data and its associated programs more portable. Team members who work with the data also become more portable since they can understand and use data with more ease. This makes the development and validation of SAS programs within a regulated environment much more efficient. This paper will present strategies on working with the new CDISC Submission Data Model version 3.0 along with other data standard strategies that are independent of CDISC. Some of the approaches include:

1. Automated evaluation of existing data structures against CDISC standards
2. Automated evaluation of existing data structures and formats among each other
3. Designing new data standards from existing data structures

The practical concepts of these techniques are demonstrated through both a manual process and tools such as %cdisc and %difftest. The meticulous review of data attributes among all the data is an important step towards achieving data standards. Automating this process makes the task less mundane, yet catches non standard differences that are not easily caught through manual verification.

Paper Number: FC04 *Audience Level*: Intermediate
WEDNESDAY, 10:00 AM-10:50 AM Room: Topaz

Submission of Analysis Datasets and Documentation: Scientific and Regulatory Perspectives

Dave Christiansen ~ Christiansen Consulting / Stephen Wilson ~ FDA

Analysis data submitted in support of new drug applications should allow statistical reviewers to identify, understand, replicate, explore and confirm the analyses performed and submitted by the sponsor. In addition to the datasets per se, the sponsor must provide clear descriptions of the creation and content of the datasets and the analyses performed.

This paper will discuss the structure and content of analysis datasets and documentation for regulatory submissions. Topics will include metadata for datasets and variables, the submission of analysis programs, the generation of analysis datasets and the concept of analysis-level metadata. The authors will discuss both the scientific and regulatory aspects of a proposed Guidance for Industry on Analysis Datasets and Documentation.

SPECIAL SESSION
WEDNESDAY, 11:00 AM-11:50 AM Room: Topaz

Panel Discussion and Open Forum on Industry Data Standards

Moderator: *Izabella Peszek ~ Merck & Co., Inc*

Panel: Dave Christiansen ~ *Christiansen Consulting*
Stephen Wilson ~ *FDA*
Edward Helton ~ *SAS Institute*
Chuck Jaffe ~ AstraZeneca
Programmer to be named later

Panel discussion on industry data standards (CDISC, AdaM, etc.), their implementation and effects on the industry, FDA and the pharmaceutical SAS programming community. Questions from the audience on these topics will be entertained for the last twenty minutes of the session.

Overview of Papers

POSTERS

CO-CHAIRS:	Brenda Bishop, EMB Statistical Solutions
	David Stokar, DataCeutics, Inc.
ON DISPLAY:	Monday & Tuesday, May 24 & 25, 8:00 AM - 5:00 PM
AUTHORS PRESENT:	Monday & Tuesday, May 24 & 25, 2:00 PM - 3:00 PM
LOCATION:	Around 2nd Floor Atrium

Paper Number: PO01 — *Audience Level*: Intermediate
MONDAY, 2:00 PM-3:00 PM — Room: Around 2nd Floor Atrium
TUESDAY, 2:00PM-3:00 PM — Room: Around 2nd Floor Atrium

Statistical Application of SAS in Method Comparison Analysis

Zaizai Lu ~ AstraZeneca Pharmaceuticals

It is common to study the agreement of measurements between two methods in clinical studies. A comparison of methods needs to be performed to estimate accuracy or errors of each method. Many techniques consisting of exploratory graphics and statistical analysis are introduced to identify the relationships between the results from two methods. All SAS codes related to the topic are also presented in this paper.

Paper Number: PO02 — *Audience Level*: All
MONDAY, 2:00 PM-3:00 PM — Room: Around 2nd
TUESDAY, 2:00PM-3:00 PM — Room: Around 2nd Floor Atrium

Be Careful When You Merge SAS Datasets

Yan Lei ~ GlaxoSmithKline

When doing SAS programming, we often need to merge two or more datasets into one. The problem that arises with merging is that when MERGE, IF-THEN statement are used within one data step, the results may be wrong or it may cause confusion.

This paper presents two examples that show clearly what happens when you use the above method to do MERGE and also provides tips to avoid these kinds of mistakes.

Paper Number: PO04 — *Audience Level*: Beginner
MONDAY, 2:00 PM-3:00 PM — Room: Around 2nd Floor Atrium
TUESDAY, 2:00PM-3:00 PM — Room: Around 2nd Floor Atrium

Testing Normality of Data using SAS®

Guangbin Peng ~ Eli Lilly and Company

Many statistical tests require data to be approximately normally distributed. Usually, the first step of data analysis is to test the normality. Also, we often test the normality of residuals after fitting a linear model to the data in order to ensure the normality assumption of the model is satisfied. SAS has offered four statistical tests that provide an easy way to test the normality. However, we should be cautious when we use these tests due to their limitations. In some cases, we may draw incorrect conclusions by only looking at the test statistics and p-values. Graphical methods are powerful in displaying distribution characteristics of the data and can serve as a useful tool in checking the normality. Combining graphic methods and statistical tests will improve our judgments on the normality of the data. In this paper, I will present these methods SAS uses by applying them to the real data from a clinical trial.

Overview of Papers
POSTERS (continued)

Paper Number: PO03 *Audience Level*: All
MONDAY, 2:00 PM-3:00 PM Room: Around 2nd Floor Atrium
TUESDAY, 2:00PM-3:00 PM Room: Around 2nd Floor Atrium

THE META-ANALYSIS (PROC MIXED) OF TWO PILOT CLINICAL STUDIES WITH A NOVEL ANTIDEPRESSANT

Lev Sverdlov ~ Innapharma, Inc.

The combined data was obtained from two double blind, placebo-controlled, clinical studies in which a novel antidepressant was used to treat a population with major depression. Inpatients (N=52) were enrolled in the first study and outpatient (N=55) – in the second study. Both studies investigated the efficacy and safety of the new drug for one or two 5-day treatment cycles. After the treatment all patients were observed during a 4-week follow-up period. The primary efficacy variable was HAMD 21. A pharmacokinetic analysis of the studies permitted the definition of the Minimum Projected Therapeutic Concentration (MPTC) and allowed for the subsequent division of the study population into three treatment subgroups: placebo subgroup (all subjects from placebo group); drug-treated subgroup 1 with drug plasma concentration (DPC) above MPTC; and drug-treated subgroup 2 with DPC below MPTC. Retrospective meta-analysis of efficacy based on individual patient data was performed in order to estimate and make inferences about a statistical difference across studies and treatment groups. A statistical model (PROC MIXED) included the following factors: the studies; the treatment groups; interaction between studies and treatment groups; and covariance from baseline. Meta-analysis showed significant differences in response between the subgroup 1 (patients with DPC>MPTC) versus both the subgroup 2 (patients with DPC<MPTC) and the placebo. There was no significant effect of the studies and interaction between studies and treatment groups. There also was consistent statistical indication of covariance for the post-treatment effect from a pre-treatment. This information prompted further clinical testing of an important novel antidepressant.

Paper Number: PO05 *Audience Level*: Beginner
MONDAY, 2:00 PM-3:00 PM Room: Around 2nd Floor Atrium
TUESDAY, 2:00PM-3:00 PM Room: Around 2nd Floor Atrium

The Paired T-Test: Does PROC MIXED Produce the Same Results as PROC TTEST?

Jack Nyberg ~ Pfizer

The paired t-test is a common technique to assess pairwise treatment differences when all treatments are given within the same experimental unit (e.g., subject). When data are paired, it is reasonable to assume that the observations within experimental unit are positively correlated. By taking advantage of that correlation, the analysis will achieve more power than by using a 2-sample t-test in which the observations within experimental unit are assumed to be independent. In SAS v.8.2, PROC TTEST can easily perform the paired t-test with the PAIRED statement. However, the data format requirements (horizontal structure) for PROC TTEST are at odds with the common data structure required to use PROC MIXED. As PROC MIXED is a standard analysis engine within the pharmaceutical industry, it is desirable that it could be used to perform a paired t-test.

In this paper, the code to perform a paired t-test in PROC MIXED is presented and is shown to produce the same estimate of the mean of the paired differences, estimate of the variance of the paired differences, degrees of freedom for the paired t-test, and resultant p-value for the paired t-test as does PROC TTEST.

Overview of Papers
POSTERS (continued)

Paper Number: PO06 — *Audience Level*: Theoretical
MONDAY, 2:00 PM-3:00 PM — Room: Around 2nd Floor Atrium
TUESDAY, 2:00PM-3:00 PM — Room: Around 2nd Floor Atrium

CLINICAL TRIAL RANDOMIZATION DESIGN: BLOCKED RANDOMIZED RESPONSE ADAPTIVE METHOD

Kwun Yee Poon ~ VA Palo Alto Health Care System/Stanford Univ / Che-Chin Lie ~ PERC

The blocked response adaptive randomization method combines the strengths of both blocked design and response adaptive methods. The SAS code of the our model divides study participants by the randomized blocked method, and then applies the response adaptive procedure to randomize study participants within each block. Through this randomization, the conflict between collective and individual ethics in clinical trial practices is balanced. We demonstrate how this SAS code can also be modified to incorporate other block randomization designs to mitigate covert biases.

Paper Number: PO07 — *Audience Level*: Intermediate
MONDAY, 2:00 PM-3:00 PM — Room: Around 2nd Floor Atrium
TUESDAY, 2:00PM-3:00 PM — Room: Around 2nd Floor Atrium

TRANSPOSITION OF VARIABLES WITH DIFFERENT FORMATS

Peter Lin ~ AstraZeneca

Sometimes with clinical trial data, several variables need to be transposed into a single variable with multiple observations. At the same time, the transposed variable needs to keep the different formatted values of the original variables. It is impossible to use PROC TRANSPOSE in one step to accomplish the transposition. This paper introduces a macro that transposes all variables (numeric or character) into one character variable while retaining the different formatted values of the original variables. If a variable does not have a format, its character value is kept in the transposed variable. Users can also define the name and the length of the transposed variable.

Paper Number: PO08 — *Audience Level*: All
MONDAY, 2:00 PM-3:00 PM — Room: Around 2nd Floor Atrium
TUESDAY, 2:00PM-3:00 PM — Room: Around 2nd Floor Atrium

UltraEdit – The Elegance and Ease of SAS Programming on a Server

Eugene Tsykalov ~ GlaxoSmithKline

This paper introduces the SAS community to UltraEdit, the Program Editor which is, in the author's opinion, best suited for SAS programming, rich on practical features and has the easiest and most intuitive work implementation with SAS files on a remote server.

UltraEdit is a Windows application, which runs on a PC and has a familiar and user-friendly interface and editing features. What makes UltraEdit unique is its ability to work with files on any remote server as easy as working with files on your local hard drive. Using a built-in FTP engine, UltraEdit allows programmers to easily open/save/reload SAS files residing on any PC-based or mainframe server and in any operating system (Unix, VMS, Windows NT, etc.).

UltraEdit combines the elegance and ease of FTP files access with practical editing features such as SAS color syntax highlighting, powerful search, column block processing, project files and many others. This shareware program (www.ultraedit.com) is capable of making a great impact on the productivity of any SAS programmer.

Overview of Papers
POSTERS (continued)

Paper Number: PO09 — *Audience Level*: Beginner
MONDAY, 2:00 PM-3:00 PM — Room: Around 2nd Floor Atrium
TUESDAY, 2:00PM-3:00 PM — Room: Around 2nd Floor Atrium

Swiss Army Knife of Listing All Filenames For Multiple Libraries

William Csont ~ H.M. Proskin & Associates, Inc.

SAS ® Software is a great tool for obtaining the same results from a variety of different programming techniques. In this paper we will examine two distinctly different methods of arriving at the same result: a report that lists individual filenames from different libraries. The older method employs SAS® macros, PROC CONTENTS and PROC SORT; the newer method employs SAS® code that utilizes a PROC FORMAT and a DATA step that accesses the SAS Dictionary Tables. The code for each method ends with a PROC TABULATE to aggregate the results in a presentation format. Each method is useful for documenting SAS® datasets and libraries.

Paper Number: PO11 — *Audience Level*: Beginner
MONDAY, 2:00 PM-3:00 PM — Room: Around 2nd Floor Atrium
TUESDAY, 2:00PM-3:00 PM — Room: Around 2nd Floor Atrium

Making Model Selection in Linear Mixed-Effects Models An Easy Process: A SAS Macro Procedure

Henry Cheng ~ MDS Pharma Services / Yuli Xie ~ MDS Pharma Services / Elizabeth Reinbolt ~ MDS Pharma Services

PROC MIXED is widely used for statistical analyses of clinical data in pharmaceutical research. Although Proc Mixed is specially designed to fit fixed-effects models, it can also model random and mixed-effects data, repeated measures, spatial data, data with heterogeneous variances and auto-correlated observations. In model selection, one wants to select the model that performs best. However, model selection can be lengthy, tedious, and time consuming. This paper presents a practical solution using a simple SAS macro procedure to make model selection an automated and efficient procedure. The same strategy can also be used for other SAS linear model selection procedures. The reader is expected to have basic statistical knowledge and be familiar with Proc Mixed and SAS macro and macro variable.

Paper Number: PO12 — *Audience Level*: Intermediate
MONDAY, 2:00 PM-3:00 PM — Room: Around 2nd Floor Atrium
TUESDAY, 2:00PM-3:00 PM — Room: Around 2nd Floor Atrium

Using Macro to Create a New Format to Transpose Differently Formatted Values

Shulin Yuan ~ Astrazeneca

PROC TRANSPOSE is a useful procedure used almost daily in SAS programming. One powerful feature is the variable created by PROC TRANSPOSE to identify the source of the values in each observation in the output data set. But this feature limits the transposition of formatted values, particularly when the variables in the input data set have mixed and/or complicated formats. This paper will demonstrate the use of a macro to create a new format to transpose differently formatted values.

Overview of Papers
POSTERS (continued)

Paper Number: PO13 — *Audience Level*: Technical
MONDAY, 2:00 PM-3:00 PM — Room: Around 2nd Floor Atrium
TUESDAY, 2:00PM-3:00 PM — Room: Around 2nd Floor Atrium

Apply Response Surface Analysis for Interaction of Dose Response Combine Treatment Drug Study

Tung-Yi (Tony) Wu ~ KOS Pharmaceuticals

When a clinical trial employs two drugs and runs expensive, a matrix design with response surface analysis become an essential statistical method to obtain overall profile, minimum effective dose and/or maximum percentage change. In this paper, the author will demonstrate one example of the SAS program using Proc RSREG combine with Proc GPLOT to obtain a response surface and one example of study design.

Paper Number: PO14 — *Audience Level*: Intermediate
MONDAY, 2:00 PM-3:00 PM — Room: Around 2nd Floor Atrium
TUESDAY, 2:00PM-3:00 PM — Room: Around 2nd Floor Atrium

Extending ODS Output by Incorporating TrellisTM Graphics from S-PLUS

Robert Treder ~ Insightful Corporation / Jagrata Minardi ~ Insightful Corporation

In the past decade many new ideas in graphical displays (Cleveland, 1993) have made their way into common usage in technical papers and reports especially in the pharmaceutical industry. In particular the TrellisTM graphics found in S-PLUS provide a level of detail in a multi-dimensional context that no other display can match in clarity nor graphical brevity. Though these displays are not available in SAS GRAPH, they may be incorporated in standard SAS ODS output by running S-PLUS alongside SAS and incorporating S-PLUS graphical output into the ODS report. This extends the possibilities of an ODS production report in a substantial way. The primary thrust of this paper is to demonstrate the simple mechanism for incorporating graphics from another non-SAS system into a standard ODS report and to describe important examples using Trellis graphics in pharmaceutical studies.

Overview of Papers

STATISTICS & PHARMACOKINETICS

CO-CHAIRS: Gajanan Bhat, Boston Scientific Corporation
Patricia Cerrito, University of Louisville
TIME: Tuesday, May 25, 1:30 PM - 5:30 PM
Wednesday, May 26, 8:00 AM - 12:00 PM
LOCATION: Diamond I

Paper Number: SP03 *Audience Level*: Intermediate
TUESDAY, 1:30 PM-1:50 PM Room: Diamond I

Automate pharmacodynamic analyses of repeated measure in clinical trials using SAS Mixed procedure

Adeline Yeo ~ Eli Lilly and Company / Grace Li ~ Eli Lilly and Company

Statistical analyses are often performed in clinical pharmacology studies to determine differences in pharmacodynamic (PD) responses of subjects due to drug administration. Variability of PD responses is also reported to support future study designs. An improvement to business process is to create a common SAS macro customized for PD analyses. With the automation of the SAS code, it will decrease our turnaround time from clinical data to specific results and speed up the process of reporting. In addition, repeated validation of similar SAS programs for PD analyses may no longer be necessary. This SAS macro is suitable for cross-over or parallel design with repeated measurements and it uses the linear mixed effects model in SAS's proc MIXED (Version 8.2 or higher).

This paper provides the macro variables needed for the operation of the SAS macro and an overview of the SAS macro. The commonly used covariance structures in PD analyses have been incorporated into the SAS macro and a discussion of it will be included.

Paper Number: SP08 *Audience Level*: Intermediate
TUESDAY, 3:00 PM-3:20 PM Room: Diamond I

Statistical Simulations for Sample Size Calculation with PROC IML

Izabella Peszek ~ Merck & Co., Inc

This paper describes the techniques for sample size calculation via statistical simulation with PROC IML. The simulation approach may turn to be the most practical way to calculate sample size in case of very complex trial designs. The concept is illustrated on a case example: validating sample size calculation program from public domain using simulation methods. The basic concepts of statistical simulation are described and corresponding IML code is provided.

Overview of Papers
STATISTICS & PHARMACOKINETICS (continued)

Paper Number: SP11 *Audience Level*: All
TUESDAY, 3:30 PM-3:50 PM Room: Diamond I

Incremental methods of imputation For Longitudinal data With Informative Missingness

Naum Khutoryansky ~ Novo Nordisk Pharmaceuticals

Incremental methods of imputation For Longitudinal data With Informative Missingness In longitudinal clinical trials, missing data are mostly related to dropouts. Some dropouts appear completely at random. The source for other dropouts is withdrawal from trials due to lack of efficacy. One of the approaches to comply with the intent-to-treat principle is the imputation of incomplete data. When the dropout process is related to the outcome process, it creates tremendous challenges in analyzing such data. No commercial software currently considers the dropout mechanisms in dealing with informative or non-random dropout. Consequently, the results are biased and misleading. This paper deals with the incremental methods of imputation applied to incomplete longitudinal data sets with informative dropouts. It is shown by simulation that the incremental methods are more precise than some other imputation methods (including the last observation carried forward method, the multiple imputation method).

Paper Number: SP05 *Audience Level*: Intermediate
TUESDAY, 4:00 PM-4:50 PM Room: Diamond I

The Analysis of Gene Markers and the Use of SAS Procedures to Examine MicroArray Data

Patricia Cerrito ~ University of Louisville

The purpose of this paper is to demonstrate SAS procedures that can be used to investigate gene marker data. Gene markers are often studied through microarray analysis where a chip containing identified genes is treated with a specific drug and then examined. The paper includes discussion of the use of SAS/Genetics and SAS/MicroArray as well as Proc KDE and Proc CALIS in SAS/Stat. Proc KDE is used to find the population distribution of gene markers for one treatment, or multiple treatments. Proc CALIS is used to develop structural equation models that can provide a means to examine the relationships between gene markers.

Paper Number: SP06 *Audience Level*: Intermediate
TUESDAY, 5:00 PM-5:20 PM Room: Diamond I

Using PROC LOESS With Sparse Data

Kenneth Liu ~ Merck / Alison Kellock ~ Merck

PROC LOESS generates “smooth” curves through data which are often used in graphs. Sparse data (=5 data points) cause computational difficulties for PROC LOESS. If extrapolation occurs when producing “scored results,” PROC LOESS issues warning messages. This paper will illustrate and provide solutions to these problems.

Overview of Papers
STATISTICS & PHARMACOKINETICS (continued)

Paper Number: SP01 *Audience Level*: Beginner
WEDNESDAY, 8:00 AM-8:20 AM Room: Diamond I

Graph the R Matrix in Linear Mixed Model

Jian Wu ~ Roche Products Australia / Peter Button

In the longitudinal studies, subjects are usually measured repeatedly over the time. Accordingly, the independence can not be assumed between the data collected within the same subject. Biased inference could result from the longitudinal data analysis without accounting for the within-subject correlation. Linear Mixed Model allows for the within-subject correlation and is a proposed model to analyse longitudinal data.

The elements in the R matrix for the Linear Mixed Model reflect the within-subject data relationship (residual covariance structure). One of the tasks for the model building is to select the most suitable residual covariance structure for the R matrix. More than 30 types of structures for the R matrix are available as an option in SAS Version 8. In addition to numerical methods, visual inspection of the various covariance structures has proved to be helpful in selecting the most suitable one for the R matrix.

Our experience with plotting the residual covariance for the R matrix is discussed in this paper. PROC MIXED, PROC GPLOT, PROC G3GRID and PROC G3D are proposed to graph the residual covariance for the R matrix. Some inappropriate methods of plotting are also discussed in the paper.

Paper Number: SP02 *Audience Level*: All
WEDNESDAY, 8:30 AM-8:50 AM Room: Diamond I

SAS Application in 2 * 2 Crossover Clinical Trial

Wuwei (Wayne) Feng ~ Eli Lilly & Company / Dong Ding ~ MedFocus, Inc.

Crossover clinical trial design attracts a lot of attention in today's drug development environment because of its unique characteristics. This paper will start with an introduction of a 2 * 2 crossover clinical trial design, then a review of advantages and disadvantages of cross over studies compared with parallel studies. Different types of SAS procedures, such as PROC TTEST, PROC GLM and PROC MIXED will be used as demonstration on how to analyze the data if response variable is continuous. McNemar test, Mainland-Gart test, Prescott test will also be reviewed if the outcome measurement is a binary variable. The paper closes with an example using simulated data.

Paper Number: SP07 *Audience Level*: Intermediate
WEDNESDAY, 9:00 AM-9:50 AM Room: Diamond I

2D-gel Proteomics in Biomarker Discovery

Mark Carpenter ~ Auburn University / Craig Rowell ~ University of Alabama at Birmingham / Coral Lamartiniere ~ University of Alabama at Birmingham

Recent advances in technology make possible the large-scale application of proteomics for biomarker discovery in cancer models and the exploration of the action mechanism of drugs. Comparative proteomics using 2-D gel technology, coupled with mass-spectrometry, is currently a widely applied tool in the discovery of unique proteins, as well as mechanisms and pathways for a variety of complex biological systems. Because 2-D gels can potentially present thousands of proteins covering dynamic ranges of expression, new approaches in data processing, database design and quantitative statistical analysis are required. In this paper we describe SAS solutions to the data processing and analysis of 2-D gel experiments, the creation of interactive and standardized databases and differential protein expression profiling. The SAS solutions involve basic DATA STEP and MACRO programming and the applications of the GLM, MIXED, TTEST, MULTTEST and GPLOT procedures. In addition, the power of ODS is exploited in the generation of derivative datasets and the standardized statistical reports.

Overview of Papers
STATISTICS & PHARMACOKINETICS (continued)

Paper Number: SP10 *Audience Level*: Advanced
WEDNESDAY, 10:00 AM-10:20 AM Room: Diamond I

Traditional repeated measures analysis versus random coefficients models Using PROC MIXED

Radhi Abdulnabi ~ Pfizer

A longitudinal data set is defined as a data set in which the response for each experimental unit is observed on two or more occasions. In the paper I will present different types of Longitudinal Data Analysis, and review the traditional approaches to longitudinal data analysis including mixed models and repeated measures analysis of variance. I will concentrate on random coefficient models showing their greater power and other advantages over traditional methods. I will try to show the similarities and differences between the repeated measures and random coefficient analyses by comparing their results using the same data set. I will use simulated data with high dropout rate to evaluate the missing data effect on these approaches. I will also present the SAS code for MIXED procedure for each method and when we have to use the time as fixed effect and when as random effect as I will try to show the differences between the time as a class variable and the time as a continuous variable.

Paper Number: SP09 *Audience Level*: Beginner
WEDNESDAY, 10:30 AM-10:50 AM Room: Diamond I

Box plots for Pharmaceutical Studies: One of the Better Ways to get a Clear Picture of Study Results

Mary Cowmeadow ~ BioStat Reports

.

Box plots , bar graphs and histograms are among the most useful plots in the pharmaceutical industry. It is possible to plot virtually anything in a clinical trial dataset with one or the other. If the data is discrete, use a bar graph or histogram; if it is continuous, use a box plot. The reason these simple plots are so appropriate is that treatment group is almost always a discrete entity. A change variable such as the difference in labs from baseline to final by treatment group is the perfect setup for box plots. For many years good box plots have not been that easy to do for general pharmaceutical users in SAS, but, with the advent of version 8, really good plots can now be found in SAS/STAT as PROC BOXPLOT.

In this paper you will explore box plots styles and options, and how the blocking variable concept solves the axis problem. Finally, footnotes are suggested to help explain and interpret the symbols used in the various styles of box plots.

Paper Number: SP04 *Audience Level*: Technical
WEDNESDAY, 11:00 AM-11:50 AM Room: Diamond I

Simultaneous Nonparametric Inference in a One-Way Layout Using the SAS System

Paul Juneau ~ Pfizer Global Research & Development

The process of discovering a novel medicine is one fraught with many unknowns. Even if a fundamental understanding of the biological systems involved in the disease process exists, the effect of a novel agent or agents on an organism may be difficult to predict or characterize. Moreover, knowledge about the corresponding measurement properties (e.g., distribution) is generally very limited. For the statistician, this setting is a perfect one to apply nonparametric statistical methods for data analysis.

A one-way layout with three or more treatments is a common study design employed in basic drug discovery research for the analysis of continuous responses. The goal of such an experimental design is often to make inferences amongst several treatments (e.g., all pair-wise comparisons, all comparisons to the control, etc.). In an attempt to address the needs of drug discovery researchers, the author has developed a set of SAS macros to perform simultaneous nonparametric inference in the one-way layout. During his presentation, he will summarize the flow and processing of his macros and their application to a few examples from drug discovery projects.

Overview of Papers

TECHNICAL TECHNIQUES

CO-CHAIRS: Syamala Kasichainula, PharmaNet, Inc.
Jack Schoemperlen, Amgen, Inc.
TIME: Monday, May 24, 8:00 AM - 5:00 PM
Tuesday, May 25, 8:00 AM - 3:00 PM
LOCATION: Diamond II

Paper Number: TT24 *Audience Level*: All
MONDAY, 8:00 AM-8:20 AM Room: Diamond II

Generating N Shell Programs for N Reports in a Clinical Trial

John Gerlach ~ Aetna

The Statistical Analysis Plan in a clinical trial identifies all the reports for a protocol, often consisting of well over one hundred Tables, Graphs, and Listings (TGL's). Assuming that the analysis programmer writes one SAS program per report, there will be as many programs that contain, hopefully, several standard features, such as a program header. Rather than employ the usual cut-n-paste process of writing SAS programs, it behooves the SAS analyst to be more efficient, as well as consistent. This paper explains a method for generating all the SAS 'shell' programs for a clinical trial.

Paper Number: TT15 *Audience Level*: All
MONDAY, 8:30 AM-8:50 AM Room: Diamond II

SAS and Microsoft Excel for Tracking and Managing Clinical Trial Data: Methods and Applications

Na Li ~ Pharmacyclics / Kathy Boussina ~ Pharmacyclics, Inc.

Microsoft Excel is widely used in the pharmaceutical industry and therefore frequently requested due to familiarity and ease of use. This paper outlines an automated data communication between Excel and SAS programs such as how to consolidate data from multiple Excel files, MS ACCESS files and other data sources, establish a data warehouse by manipulating and integrating information in SAS, and deliver information in Excel format applying Excel templates with reusable macros or by dynamically creating multiple Excel worksheets. Several SAS methods are explored including DATA step, SAS/ACCESS (PROC ACCESS, SQL Procedure, PROC DBLOAD, PROC IMPORT/EXPORT), DDE (Dynamic Data Exchange), and ODS (Output Delivery System). The pros and cons of each method are also summarized. Furthermore, the scheduling of SAS executions on a daily basis is addressed. The application of these methods is further detailed, illustrating how they can help in the management of clinical trials when delivering information to Clinical Research Associates, and Medical and Safety Monitors. The powerful data warehousing techniques and information delivery methods of the SAS system are utilized to integrate multiple data sources and deliver consolidated Microsoft Excel worksheets to medical professionals.

Overview of Papers
TECHNICAL TECHNIQUES (continued)

Paper Number: TT18 *Audience Level*: Technical
MONDAY, 9:00 AM-9:20 AM Room: Diamond II

An Introduction to SAS Applications of the Windows Scripting Host

Stephen Hunt ~ PRA International / Tracy Sherman ~ PRA International / Brian Fairfield-Carter ~ PRA International

Analysis Programmers in the Pharmaceutical industry are often confronted with a variety of requirements that are peripheral to the more common tasks of data manipulation and summarization. These include supporting multiple output file types (.doc, .rtf, .ps, .pdf), making operating system calls (for example, to retrieve file attribute information), and carrying out file-handling operations. These tasks are often satisfied through more obscure components of Base SAS (DDE, System functions, 'pipe' filename access, x/DOS commands), and through code-writing macros (custom output 'drivers'), that allow SAS interaction with other applications.

This paper introduces the Windows Scripting Host (WSH) as a simple, powerful, and comprehensive tool for meeting all of the requirements listed above. The WSH allows VBScript, Jscript and Perl programs to run on the host operating system as stand-alone applications, and thus gives SAS programmers not only an environment in which to program in COM (the 'Component Object Model'), but also to develop dynamic applications. This paper focuses on the fundamentals and practical application of VBScript for running Operating System and File System commands and for instantiating and controlling Windows applications, and on methods for integrating WSH applications with a SAS session. The final section introduces the ActiveX Data Object for writing data between SAS, Excel, and Access.

Paper Number: TT13 *Audience Level*: Intermediate
MONDAY, 9:30 AM-9:50 AM Room: Diamond II

USING THE SAS WINDOWING ENVIRONMENT FOR PROGRAM DEVELOPMENT

Wei Cheng ~ Isis Pharmaceuticals, Inc.

The SAS System Windowing Environment provides many new features to assist SAS programmers in developing their SAS programs and applications. This paper gives examples of customizing the SAS Windowing Environment to build a program development environment. Topics include building a help page to find answers to SAS questions, building a toolbar icon to scan the Log window, setting up Favorite Folders in the SAS Explorer, setting up File Shortcuts in the SAS Explorer, and setting up colors for different windows.

Paper Number: TT02 *Audience Level*: All
MONDAY, 10:00 AM-10:20 AM Room: Diamond II

Creating Word Tables using PROC REPORT and ODS RTF

Carey Smoak ~ Roche Molecular Systems, Inc.

With the introduction of the ODS RTF destination, programmers now have the ability to create Word tables using SAS. This paper focuses specifically on the use of the ODS RTF destination with PROC REPORT. Demographic data (age, race, sex, height and weight) from a research study is used to illustrate how to create Word tables. Some basic methods for generating Word tables are demonstrated. Several problems are illustrated and various solutions are presented. Some common tasks that I illustrate are: 1) creating headers, titles and footnotes, 2) page numbering (Page X of Y), 3) the bodytitle option and its effect on headers, titles, footnotes and page numbering, and 4) breaking a page. The methods presented in this paper should help you get started with the task of creating Word tables with ODS RTF and PROC REPORT. The Version 8.2 of SAS on a Windows platform was used.

Overview of Papers
TECHNICAL TECHNIQUES (continued)

Paper Number: TT11 *Audience Level*: Intermediate
MONDAY, 10:30 AM-10:50 AM Room: Diamond II

Producing Data-Driven Plots for Microsoft Word Using SAS/GRAPH and the SAS Macro Language

Robert Graebner ~ Quintiles, Inc.

SAS/GRAPH is a powerful tool that provides a broad range of graphing capabilities. When combined with the power of the SAS macro language, SAS/GRAPH allows users to develop flexible applications that can automate the production of graphs. The ability to output graphs into the most common graphic file formats allows publication quality graphs to be easily created and imported into the commonly used personal computer applications. This paper presents the basics of using the SAS macro language to automate the process of producing graphs and importing them into Microsoft Word. Four basic uses of the macro language are covered; automatically generating plots for all defined subsets of data, tailoring each plot to the data in a particular subset, outputting the graphs to computer graphics metafiles (CGM) for importing into Microsoft Word and generating Visual Basic source code to insert the figures into a Word document and resize them. Examples from a clinical trial study in which over 3,000 graphs were produced are used to illustrate the key concepts involved. This paper is intended for users with basic SAS/GRAPH experience who are interested in expanding their graphing capabilities through the use of the SAS macro language.

Paper Number: TT07 *Audience Level*: Theoretical
MONDAY, 11:00 AM-11:20 AM Room: Diamond II

Instant SAS Applications With VBScript, Jscript, and dHTML

Brian Fairfield-Carter ~ PRA International / Tracy Sherman ~ PRA International / Stephen Hunt ~ PRA International

For a programmer with limited experience outside of Base SAS, the objective of setting up an application that runs SAS as a 'back-end', and that works with SAS source code outside of the SAS system, can be somewhat daunting. Naturally, the first thing you'd ask would be 'what is the simplest, cheapest environment that I can get started with?' A good answer seems to lie in the combination of Windows Scripting technologies (VBScript and Jscript), dynamic HTML, and SAS integration with the Windows Component Object Model (COM). HTML provides a very simple environment in which to create a graphical user interface, which can be made dynamic (as in, responsive to end user actions) by placing VBScript and Jscript components between <SCRIPT> and </SCRIPT> tags. The script components can, in turn, instantiate SAS sessions, run SAS code, and retrieve and display SAS output and data. This environment is attractive as a starting point for applications development because it is simple and cheap: it requires no additional software installation (all the necessary language interpreter/compiler facilities already exist in the Windows operating system and web browser), languages (particularly VBScript and HTML) are very easy to learn, and components can be easily borrowed from other HTML pages.

Paper Number: TT12 *Audience Level*: Intermediate
MONDAY, 11:30 AM-11:50 AM Room: Diamond II

MORE CUSTOMIZATION?: CREATING SYMBOLS IN RTF FILES USING ODS

Richard Rowell ~ Connetics Corporation / Jim Lenihan ~ ASG-Inc

If you are using SAS ODS to produce RTF files you may run into problems if you need to put symbols in your tables or listings. This paper will show you how to display various symbols in your output. Additionally, it will show some of the problems the authors have had in creating certain symbols and how you can "troubleshoot" if you have difficulty in putting out your symbol.

Overview of Papers
TECHNICAL TECHNIQUES (continued)

Paper Number: TT01 *Audience Level*: All
MONDAY, 1:30 PM-1:50 PM Room: Diamond II

Different decimal places for different lab tests.

Cecilia Casas-Mauldin ~ PPD Development

This paper will show how to present the value of one variable in multiple formats. The representation of a variable can depend on the value of a second variable or in the value of the observation itself. In this case, the representation of the results of a laboratory test will depend on the name of the laboratory test. The format used for each laboratory test is independent from other laboratory tests.

Each laboratory test can have its own appropriate number of significant decimal places and the number can vary from test to test. This method guarantees that significant results are not truncated, and also, only significant decimals are presented.

Through this paper, a series of examples will show how to use the PUTN function with two, three or four parameters, and the difference between rounding or not rounding before formatting.

Paper Number: TT06 *Audience Level*: All
MONDAY, 2:00 PM-2:50 PM Room: Diamond II

'LAG with a WHERE' and other DATA Step Stories

Neil Howard ~ i3 Data Services

This presentation illustrates the importance of understanding the internals of DATA step processing, particularly when you invoke mysterious functions like LAG, code RETAIN statements, rely on WHERE constructs, and start combining various features of the DATA step language. This anthology includes such spellbinding stories as "LAG with a WHERE", "When RETAIN doesn't retain", "To LAG or to LEAD", "Don't Order My Variables Around", "A DIFferent LAG", and "The Case of the Missing Values".

Paper Number: TT10 *Audience Level*: Intermediate
MONDAY, 3:00 PM-3:20 PM Room: Diamond II

The FORMAT procedure – more than just a VALUE statement

Lawrence Heaton-Wright ~ Quintiles Limited

The FORMAT procedure is most frequently used to define formats for variables. However, there is also extended functionality available within the format procedure which allows users to define informats, picture formats, create data sets of formats, use data sets (also known as control data sets) to create formats and to display the formats catalog information.

This paper aims to show users how informats and picture formats work, how to create and use control data sets and how to display the formats catalog. The audience for this paper does not have to be of a high technical level but should be familiar with the workings of base SAS. The contents are not restricted to one particular operating system.

Overview of Papers
TECHNICAL TECHNIQUES (continued)

Paper Number: TT22 *Audience Level*: Advanced
MONDAY, 3:30 PM-3:50 PM Room: Diamond II

MULTILABEL - A useful addition to the FORMAT procedure

Venkatesan Chakravarthy ~ Independent Consultant

PROC FORMAT is almost an indispensable tool for the SAS programmer. One of its limitations, until recently, was that it could accept only mutually exclusive categories. The MULTILABEL option introduced in version 8.2 overcomes this limitation and has some powerful applications in generating tables. This paper illustrates MULTILABEL with an example to summarize by age categories where some age values are mapped to more than one category. The paper also addresses some of the limitations of this new option. A bonus feature of this paper is the coverage given to some useful enhancements to PROC MEANS that are especially relevant to summary table generation in the Pharmaceutical Industry.

Paper Number: TT23 *Audience Level*: Intermediate
MONDAY, 4:00 PM-4:20 PM Room: Diamond II

Dualing Arrays: Reducing DATA Step Passes with Array Look-Ups

Nancy Brucken ~ STATPROBE, Inc. / Lisa Reeves ~ STATPROBE, Inc.

As part of an analysis of the effect of treatment group and entry order (subject number) on change in certain laboratory values, we were requested to provide pairwise comparisons of treatment effects by subject number. With the Output Delivery System (ODS), we were able to retrieve the required p-values, and store them in a SAS dataset. The problem then became one of extracting only the pre-planned comparisons. Using the DATA step in a non-traditional way and building look-up tables with arrays allowed us to identify the LSMEAN numbers for the desired comparisons and then pull off the p-values calculated for those comparisons in a single DATA step, with only one pass through each of the associated datasets.

Paper Number: TT17 *Audience Level*: Intermediate
MONDAY, 4:30 PM-4:50 PM Room: Diamond II

Creating PDF files for SAS Programs

Eric Zhang ~ Merck & Co., Inc.

Documents produced in the Portable Document Format (“PDF”) are becoming popular among pharmaceutical companies. SAS programmers might wish to archive SAS programs after projects are finished in order to protect them from unexpected modifications. Creating PDF files of SAS programs is one of the best ways to achieve this goal. This is due to the fact that Acrobat Reader is available free on the internet and PDF files can not be easily modified. There are many tools available to convert SAS programs to PDF files and/or build hyper-links between the PDF file, “Table of Contents”, and individual PDF files. However, these tools are developed outside SAS and might not be portable across operation systems. It’s also hard to modify existing features of these tools for the specific requirements we need. In order to avoid the inconvenience of using existing tools, a SAS macro, called PGM2PDF, was developed to have these similar features without any third-party software support. In addition, it’s easy to use and modify. This paper describes the capabilities, features, structure, and usage of this macro.

Overview of Papers
TECHNICAL TECHNIQUES (continued)

Paper Number: TT21 *Audience Level*: Advanced
TUESDAY, 8:00 AM-8:20 AM Room: Diamond II

One Frequency Macro for All Frequency Tables

Sam Mao ~ Quintiles, Inc / Jim Edgington ~ Quintiles

Frequency tables are one of the major components of safety analysis of clinical trial; those frequency tables include various AE tables, concomitant medication summary tables, and subject counting tables of study participation. While these tables vary in forms, one common characteristic is that they all calculate and present frequency, from one level frequency to three levels of frequency in a single table. Following modular design principal, one frequency macro has been developed to do the common task. Using this macro as backbone together with other macros (modules), any frequency tables can be generated. In this paper, the design and implementation of the macro is discussed, examples of using the macro are demonstrated.

Paper Number: TT14 *Audience Level*: Intermediate
TUESDAY, 8:30 AM-8:50 AM Room: Diamond II

Programming Tips and Examples for your Toolkit

John Morrill ~ Quintiles, Inc / Kristi Wiser ~ Quintiles, Inc.

This paper describes several independent ideas that perhaps should be in the toolkit of intermediate level programmers. We cover methods of supplying missing summarization levels using SAS 8.2 PROC MEANS options and the DATA step UPDATE statement. We examine a nifty example of adverse event reporting by month. We suggest an alternative to standard if-then-else logic, the DATA step SELECT statement. We review Proc Tabulate and a macro to obtain a quick glance at all data sets in a directory. We also present some thoughts on formats and a few other ideas along the way.

Paper Number: TT19 *Audience Level*: Advanced
TUESDAY, 9:00 AM-9:50 AM Room: Diamond II

Undocumented and Hard-to-find PROC SQL Features

Kirk Lafler ~ Software Intelligence Corporation

The SQL Procedure contains many powerful and elegant language features for advanced SQL users. This paper presents PROC SQL topics that will help programmers unlock the many hidden features, options, and other hard-to-find gems found in the SQL universe. Topics include CASE logic; the COALESCE function; SQL statement options _METHOD, _TREE, and other useful options; dictionary tables; automatic macro variables; and performance issues.

Paper Number: TT05 *Audience Level*: Intermediate
TUESDAY, 10:00 AM-10:20 AM Room: Diamond II

Controlling Graph Size: Building Thumbnails and GIF Files Using SAS/GRAPH

Art Carpenter ~ CA Occidental Consultants / Richard Smith ~ Data Explorations

GIF is one of the primary file forms used for graphs that are to be displayed on the web. However because of differences in browsers and the need to control the size of the graphic display, the standard GIF device that is defined with SAS/GRAPH is not always appropriate. This paper discusses alternative methods that can be used to create GIF files that will display graphs of various sizes.

Since secondary considerations as to graph contents, options, and control alternatives become necessary when the size of a graphic display is reduced, these issues are also presented.

Overview of Papers
TECHNICAL TECHNIQUES (continued)

Paper Number: TT09 *Audience Level*: All
TUESDAY, 10:30 AM-10:50 AM Room: Diamond II

Is the Legend in your SAS/GRAPH Output Telling the Right Story?

Justina Flavin ~ Pfizer Global Research & Development / Arthur Carpenter ~ California Occidental Consultants

When generating plots using the GPLOT procedure, a third variable, such as a treatment or dose group, is often used as a grouping variable. GPLOT automatically uses the values of this variable in the LEGEND, but very often the designers of the plots will want particular symbols, line types, and/or colors to be specified for the display of specific values of this variable. Of course consistency is important, and if the program is subsequently run on a subset of the original data, some values of this discrete variable may be absent in the subset. This can result in an inconsistent representation (symbol, line type, color) of a particular value of the variable in a series of plots.

This paper will present macro and non-macro solutions for producing plots with consistent attributes for each distinct value of the grouping variable, showing how to maintain this consistency regardless of the subset used. The paper will also explain how to generate a legend which will either display all values that exist in the data or only those that appear in the subset.

Paper Number: TT08 *Audience Level*: Intermediate
TUESDAY, 11:00 AM-11:20 AM Room: Diamond II

Processing Large Lists of Parameters and Variables With Macro Techniques

Eugene Tsykalov ~ GlaxoSmithKline

Have you ever been in a situation when you need to process a large number of variables and/or parameters? For example, validating correctness of units and presence of parameters in a large normalized (vertical) dataset. Add to it a cross checking of variables of different types from a normalized dataset and you've got a long code to go.

This paper presents an elegant solution to replace endless "if then", "where" and "select" statements and avoid using names of variables/parameters again and again. Use of SAS arrays and Macro language is demonstrated to create and automatically process a macro lists of variables and/or parameters of any length. These lists can be automatically divided into lists of character and numeric variables/parameters for separate processing.

Practical examples include manual transpose of normalized datasets and validation of Laboratory parameters in a clinical trial data.

Paper Number: TT20 *Audience Level*: Intermediate
TUESDAY, 11:30 AM-11:50 AM Room: Diamond II

Generating Least Square Means, Standard Error, Observed Mean, Standard Deviation and Confidence Intervals for Treatment Differences Using Proc Mixed

Richann Watson ~ Kendle International Inc.

Have you ever wanted to calculate the confidence intervals for treatment differences or calculate the least square means using a mixed model but can’t always recall the correct options or layout of the model for PROC MIXED? If so, the macro presented in this paper will appeal to you. The DOMIXED macro allows for the calculation of least square means, standard error, observed mean, standard deviation and confidence intervals for treatment difference. The macro will also calculate p-values.

Overview of Papers
TECHNICAL TECHNIQUES (continued)

Paper Number: TT03 *Audience Level*: Intermediate
TUESDAY, 2:30 PM-2:50 PM Room: Diamond II

%Cox_PH – An enhancement of SAS PROC PHREG

Hong Zhang ~ Merck & Co., Inc.

SAS PROC PHREG is widely used to model time to event data via Cox regression; however, without the familiar CLASS and CONTRAST statements available in other SAS procedures, extra programming effort is required to handle categorical variable(s) or interactions between two covariates in the model, or to test the significance of a linear combination of the regression parameters. The Survival Analysis Macro (SAM) committee at Merck reduced the limitations of PROC PHREG through the development of a %Cox_PH macro. The %Cox_PH macro provides the user flexibility to specify categorical variables or interactions between two covariates in the Cox regression model, as well as the capability to make pair-wise comparisons, possibly with different reference groups, among different levels of a multi-leveled categorical variable. This is particularly useful when comparisons between three or more treatment groups are required. In addition, the macro generates document-ready Microsoft Word or RTF tables that can be used for regulatory reporting. This paper describes the basic features and structure of the %Cox_PH macro.

Paper Number: TT16 *Audience Level*: Beginner
TUESDAY, 3:00 PM-3:20 PM Room: Diamond II

An Experiment with Experimental Proc RobustReg

Shuang Lu ~ Merck & Co., Inc.

Proc RobustReg is an experimental procedure in SAS/STAT version 9. It implements the most commonly used robust regression techniques, including M (Maximum likelihood-like) estimation, LTS estimation, S estimation and MM estimation. A macro developed by the Merck Research Laboratories, Merck & Co., Inc. also carries out M estimation of General Linear Models. This paper provides a comparison of the output from an experimental data test and suggests possible improvements in Proc RobustReg, based on the author's experience with the in-house macro.

Overview of Papers

TUTORIALS

CO-CHAIRS: Kent Letourneau, PRA International
Alfredo Rojas, Merck & Co., Inc.
TIME: Monday, May 24, 8:00 AM - 3:00 PM
LOCATION: Topaz

Paper Number: TU03 *Audience Level*: Beginner
MONDAY, 8:00 AM-8:20 AM Room: Topaz

Clinical Trials Terminology for SAS Programmers

Sy Truong ~ Meta-Xceed, Inc.

The drug development process is a clinical process that has its own language. It is not required that SAS programmers function as an MD or regulatory expert, but working knowledge of the terminology is important to be effective. This paper will walk through the drug development process from discovery through Phase IV. It will explain a wide range of acronyms such as IND, NDA, GCP and MedDRA. It will also describe some of the terminologies used within the process of clinical trials as a drug is developed and submitted to the FDA. This will give SAS programmers a larger perspective or context to the work of analyzing and reporting clinical trials data.

Paper Number: TU04 *Audience Level*: Intermediate
MONDAY, 8:30 AM-8:50 AM Room: Topaz

Using XML to Create Excel Spreadsheet in SAS :- A step by Step guide

Raj Juneja ~ Makro Technologies

Creating MS Excel Spreadsheets with SAS DDE has a variety of shortcomings. Starting with SAS 8.2, SAS XML support has a solution to these shortcomings. This paper covers how XML and Excel interface with each other, how SAS can be used to populate the Excel Spreadsheets with the help of XML Spreadsheet Language(XMLSS). It also gives examples of XMLSS tags taht can be used in SAS with generic XML to format Excel Spreadsheet. This paper assumes the readers have basic SAS and XML knowledge.

Paper Number: TU01 *Audience Level*: All
MONDAY, 9:00 AM-9:50 AM Room: Topaz

How SAS Thinks

Neil Howard ~ i3 Data Services

The DATA step is the most powerful tool in the SAS system. Understanding the internals of DATA step processing, what is happening and why, is crucial in mastering code and output. Concepts covered:

- Logical Program Data Vector (LPDV or PDV),
- automatic SAS variables and how are they used,
- the importance of understanding the internals of DATA step processing,
- what happens at program compile time,
- what's happening at execution time,
- how variable attributes are captured and stored,
- handling data defaults, conversions, and missing values.

This paper focuses on techniques that capitalize on the power of the DATA step and working with (and around) the default actions. By understanding DATA step processing, you can debug your programs and interpret your results with confidence.

Overview of Papers
TUTORIALS (continued)

Paper Number: TU06 *Audience Level*: All
MONDAY, 10:00 AM-10:20 AM Room: Topaz

A Visual Introduction to SQL Joins

Kirk Lafler ~ Software Intelligence Corporation / Charles Shipp ~ Shipp Consulting

Real systems rarely store all their data in one large table. To do so would require maintaining several duplicate copies of the same values and could threaten the integrity of the data. Instead, IT departments everywhere almost always divide their data among several different tables. Because of this, a method is needed to simultaneously access two or more tables to help answer the interesting questions about our data. This paper visually illustrates how a join process works and then shows how a PROC SQL query is constructed so some or all of the specified tables contents can be brought together.

Paper Number: TU07 *Audience Level*: Beginner
MONDAY, 10:30 AM-10:50 AM Room: Topaz

How to Develop User Friendly Macros

Sy Truong ~ Meta-Xceed, Inc.

SAS macros help automate tasks that are done repeatedly. However, if they are not easy to use or debug, they may never be used. This paper will describe approaches to developing and maintaining SAS macros that are easy to use. Some of the topics covered include:

- effective documentation of macro header
- portable code for use with different OS
- error and warning message handling
- paper and online documentation
- use of nested macros and nested macro variables
- keeping macros simple for debugging

A little effort can go a long way towards creating a successful SAS macro. This paper will present tips and techniques that are not always obvious. Besides getting the resulting numbers to the user, a user friendly macro can enhance the entire experience.

Paper Number: TU02 *Audience Level*: Beginner
MONDAY, 11:00 AM-11:50 AM Room: Topaz

Building and Using User Defined Formats

Art Carpenter ~ CA Occidental Consultants

Formats are powerful tools within the SAS System. They can be used to change how information is brought into SAS, how it is displayed, and can even be used to reshape the data itself. The Base SAS product comes with a great many predefined formats and it is even possible for you to create your own specialized formats.

This paper will very briefly review the use of formats in general and will then cover a number of aspects dealing with user generated formats. Since formats themselves have a number of uses that are not at first apparent to the new user, we will also look at some of the broader application of formats.

Overview of Papers
TUTORIALS (continued)

Paper Number: TU05 *Audience Level*: Intermediate
MONDAY, 1:30 PM-1:50 PM Room: Topaz

So you want to write a PharmaSUG paper? That paper about writing a paper

Dianne Louise Rhodes ~ Westat

You just attended your first PharmaSUG conference, and you come out of one really good session thinking, "Hey, I could write a paper like that about My Project!" That's what happened to me after I attended my first NESUG in 1996. And it has happened to others, just recently, in fact, a colleague asked me where he could find "That Paper about Writing a Paper." We couldn't find anything providing specific guidelines on how to approach the process. This paper will introduce the paper writing process: choosing a topic, researching it, developing an outline, and actually writing the paper. I'll also walk through developing a presentation from an outline created in PowerPoint. That's the best way I've found to ensure that the presentation and paper are in sync. I'll go over tips on making an effective presentation. Finally, in the reference section I'll provide other resources.