Using SAS® in Clinical Research

Greg Nelson, ThotWave Technologies, LLC.
Outline

- Introduction and Overview
- SAS – 30 years of evolution
- The SAS 9 Platform
- Clinical Research Imperatives
- Summary and Conclusions
<table>
<thead>
<tr>
<th>Overview</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAS 9</td>
</tr>
<tr>
<td>Examples</td>
</tr>
<tr>
<td>Choosing the right environment</td>
</tr>
<tr>
<td>Summary</td>
</tr>
</tbody>
</table>
Overview

• Major computing trends in Pharma
• The clinical research information value chain
• Standards that matter when it comes to DI (CDISC, ICH, HL7)
• Pieces and parts of SAS
• The roles for SAS (and SAS professionals) in pharma
BI and DI in Pharma

a 13 to 15 percent annual increase in business intelligence and data warehousing spending
• With enhanced decision-making capabilities, and the ability to respond to health authority requests and inquiries in a timelier manner, the concerted organizational effort leading up to NDA review/approval is being accelerated by up to five days – leading to earlier revenue recognition. Revenue is being estimated at $1,000,000 per day, each day the review or approval is accelerated.

• With enhanced decision-making capabilities along with the ability to utilize all drug discovery and development data in a consistent, accurate and trusted manner, there is potential to “kill” at least one more drug per year earlier in the cycle; and/or prevent one from entering a clinical program. Note: 9,999 out of 10,000 drug candidates fail to make it to the market; and for the one that makes it, the research and development costs is almost $750 million. The cost avoidance is estimated to be in the millions.
• Business intelligence (BI) is a broad category of application programs and technologies for gathering, storing, analyzing, and providing access to data to help enterprise users make better business decisions.

• BI applications include the activities of decision support, query and reporting, online analytical processing (OLAP), statistical analysis, forecasting, and data mining.

• a set of concepts and methods to improve business decision making by using fact-based support systems
• Traditionally what we think of as “data management”
• Involves combining data from lots of different places, formats and structures
• Used to create a single, credible version of the truth
• Data Integration is about beating data into submission – getting the data right before we can use it.
Corporate Information Factory

[Diagram of Information Factory: A Conceptual Architecture For Business Intelligence]

- Library & ToolBox
- Workbench
- Integration/Transformation (Data Re-engineering)
- Data Warehouse
- Data Delivery
- Operational Data Store
- Meta Data Management
- Support Data
- Transactional Systems Operational
- External Data
- Internal Data
- Real-Time Data
- Other Results
- API
- XML
- Oracle
- SAS

[Other elements labeled: Systems Management, Data Acquisition Management, Operations & Administrations, Service Management, Change Management]
How Drugs Are Used

- Prevention
- Cure
- Treatment
- Diagnosis
- Promotion
Understanding the Drug Development Life Cycle

Pharma Value Chain

Pre-Clinical Development
- Sample Management
- Tissue Bank
- Assay Results Management
- Formulation Development
- Pre-Clinical Safety Data Warehouse

Clinical Development
- Electronic Data Capture
- Clinical Trial Management
- Clinical Trials Supplies Management
- Clinical Data Management
- Regulatory Submissions
- Pharmacovigilance

Manufacturing/Operations
- SAP/Oracle Implementation & Maintenance
- CAPA Management
- Product Complaint System
- Warehouse Management System (WMS)
- Supply Chain Optimization
- Logistics Planning

Sales & Marketing
- Enterprise Customer Master
- Territory Alignment
- Smart Routing
- CRM/SFA Implementation & Maintenance
- Business Analytics
- Report Shop

Regulatory Compliance
- 21 CFR Part 11 Assessment and Remediation Services
- Computer System Validation Services
- Regulatory Compliance Tracking System
Traditional Clinical Research

Chemical Synthesis → Pre-Clinical Testing → Phase I → Phase II → Phase III

IND → NDA → Approval → Product Launch

8 - 12 years
$250 - $350 million

20 year patent protection
Pharmaceutical companies have made a tremendous investment...

- Only 1 in 10,000 of the compounds synthesized ever reach market
- An NDA is approved for 20% of the drugs which had an IND filed
- By the time a drug is approved, generally only 8 years remain on the patent
- Only 25% of all drugs that are marketed achieve profits that exceed the development costs

Hence, the importance of:

- efficient processes
- meeting or exceeding timelines
Core Functions – Targets for Improvement

- Protocol Design and Study Start-Up
- Patient and Investigator Recruitment
- Clinical Trial Management
- Clinical Data Management
- Data Analysis
- Clinical Supplies
- Regulatory and Safety
## Technology Map for Clinical Trials

<table>
<thead>
<tr>
<th>Protocol Design</th>
<th>Recruitment</th>
<th>Trial Management</th>
<th>Clinical Data Mgt.</th>
<th>Data Analysis</th>
<th>Clinical Supplies</th>
<th>Regulatory &amp; Safety</th>
</tr>
</thead>
</table>

### Key Application Functionality
- Portals
- Collaboration/KM
- Decision Support
- Electronic Data Capture
- Advanced Data & Database Mgmt.
- Visualization
- Workflow Mgmt.
- Statistical Analysis & Reporting
- Drug Supply & Tracking
- Document Mgmt.
- Project & Portfolio Mgmt.

### Integration & Aggregation
- Data Mining
- Data Warehousing
- Enterprise Application Integration
- Stds. Repository
- Enterprise Vocabulary
- Object Models
- Electronic Signatures
- Web Services

### Key Infrastructure

© Copyright 2006 ThotWave Technologies, LLC
Technology Challenges

- Validation
- Manageability
- Scalability (Performance)
- Security
- Data (and process) governance
- Auditability
• **compliance** with 21 CFR part 11: security, audit trail, version control
  - **Validation** - accuracy, reliability, consistent intended performance and the means to discern invalid or altered records
  - **Restriction of system access** to only authorized individuals
  - **Secure, computer-generated, time-stamped audit trails** to record operator entries and action to create, modify and delete electronic records – retained for required period and available for agency review and copying
  - **Operational system checks** to enforce permitted sequencing of steps and events as appropriate
  - **Authority checks** for use, e-signature, access of input and output device, altering a record and performing operation at hand

• **PDUFA III** - “Reviewable Units”
• **CDISC**
Related Standards

• ICH
  – The International Conference on Harmonization has been compiling a series of guidelines for the preparation, design, conduct, and reporting of clinical trials with an aim to harmonize the interpretation and application of technical guidelines and requirements for product registration.

• CDISC
  – Operational Data Model (ODM) - operational support of data collection
  – Study Data Tabulation Model (SDTM) – data tabulation data sets
  – Case Report Tabulation Data Definition Specification (CRTDDS - aka define.xml)
  – Laboratory Data Model (Lab)
  – Standard for Exchange of Non-clinical Data (SEND)
  – BRIDG - Protocol Representation
  – Analysis Data Model (ADaM) – analysis data structures

  – And others.. LAB, SEND
Scope of this paper

How can SAS 9 be used in Clinical Research?
Clinical Research Objectives

- To provide clinical information that is useful in the making of business and economic decisions.
- To provide understandable information which will aid stakeholders in predicting the safety and efficacy of a compound.

Qualitative Characteristics of Clinical Reporting

- Relevance
  - Timeliness
  - Predictive Value
  - Feedback Value
- Reliability
  - Verifiability
  - Neutrality
  - Representational faithfulness
- Comparability (including Consistency)
Challenges in Repeatability

- Define additional data points as the study's hypothesis is refined
- Clinical QA may revise core trial processes such as adverse event adjudication
- Trial managers sometimes devise new metrics that better reveal trial progress
- FDA often requests additional analyses that require new derived variables
- Significant change occurs between trials as well, as study teams invent new case report form (CRF) elements and technological advances introduce new ways to measure product safety and efficacy
Pervasive Impact Change Has
• Compliance/Validation
• Change control
• Auditing
• Security
Overview

**SAS 9**

Examples

Choosing the right environment

Summary
• Overview of the Platform (IVC)
• Tools that make up the platform
• What specifically is BI (in SAS 9)
• Related Aspects (DI and Analytics)
What is the SAS Information Value Chain?

- It is a marketing message from SAS
SAS 9 Components

Plan

ETL

Intelligent Storage

Business Intelligence

Analytic Intelligence

SAS DI Studio
SAS Data Surveyors
SAS Data Quality Server
SAS OLAP Administrator
SAS OLAP Server (SAS Application Server)
SAS Access Products

SAS BASE (as a transformation engine)

SPDE
SAS Scalable Performance Data Server

SAS AppDev Studio
SAS Information Delivery Portal

SAS Information Map Studio
SAS Web Report Studio
SAS Enterprise Guide
SAS Office Integration

SAS Management Console is part of the Manageability cornerstone
SAS Workspace Server and SAS Stored Process Server are used by all the Clients
• Traditional SAS using DM or, preferably, EG
  – Evolutionary. Works just like the ‘old’ SAS, but with some enhanced functionality (i.e. new procedures, functions, SPDE, etc.).

• The new client/server paradigm *(metadata-managed platform)*
  – Revolutionary! If you really want to take advantage of the new power of SAS, you will have to make the leap.
• Keys that make this more challenging in Clinical:
  – Regulatory
  – Reuse (repeatability)
  – Flexibility
  – Security and Auditability

• XML
  – Movement away from datasets to self-describing data constructs
  – Can we use these as “information”
Overview

SAS 9

Examples

Choosing the right environment

Summary
Examples: Non-clinical

• Non-Clinical Uses
  – Sales and Marketing
  – Manufacturing
  – Finance
  – Human Resources
  – Information Services
  – Executive and Portfolio management
Examples: Clinical

- Clinical Research
  - Pre-clinical Research
  - Clinical
  - Stat/Programming
  - Supporting other groups
    - Data Management (patient profiles)
    - Medical writing
    - Finance
    - Project management
    - Patient Registries & Post marketing surveillance
Sales / Marketing

Capability:
• Single version of the truth for customer
• Data mining supports the business

Users:
• Business analysts
• Case managers
• Account managers

Data:
• SAS data warehouse
• Internal Sales planning data
• IMS Sales data

ROI:
• Score defined segments “on the fly”
• Eliminates need to score entire database
• Use only the “freshest” up-to-date data
• Reduces manual intervention and error
• Accelerates the market cycle
• Increases likelihood of reaching and influencing customers and prospects with the offer at the right time
• Improves campaign results and lowers costs

Positively influence the behavior of customers and prospects through a stream of pertinent communications
Patient Registry

Capability:
- Worlds largest patient registry
- Access to research analytics
- Publication quality data available

Users:
- Clinicians
- Sales staff

Data:
- Clinical data management system

ROI:
- Afforded sales staff access to doctors
- Provided much needed clinical research findings
- Remove need for distribution of data/reports for over 1M patients at 1500 hospitals
Contract Research Organization

**Capability:**
- Modern SAS infrastructure
- Repository for all clinical information
- Assets
- Developer and end user tool

**Users:**
- SAS Programmers & Statisticians
- Medical Writers
- Data Management

**Data:**
- Existing Oracle-based data management system
- Import various unstructured data

**ROI:**
- Significantly reduce time/manpower to find data, prove that process was followed
- Provide infrastructure for global team collaboration
- Leverage a validated environment
**Organizational Planning**

**Capability:**
- Measurement based leadership
- Repository for all measurement data
- Decision support tool

**Users:**
- Executives and managers in Fortune 1000 companies
- Business consultants

**Data:**
- Survey data about customer, employee satisfaction
- Accounting data (financial metrics)

**ROI:**
- Provide a single platform for running your business
- Combined historical data with predictive modeling
- Leverage a validated environment
<table>
<thead>
<tr>
<th>Overview</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SAS 9</td>
<td></td>
</tr>
<tr>
<td>Examples</td>
<td></td>
</tr>
<tr>
<td><strong>Choosing the right environment</strong></td>
<td></td>
</tr>
<tr>
<td>Summary</td>
<td></td>
</tr>
</tbody>
</table>
• The SAS Intelligence Platform consists of a multiple tier environment that is typically represented by the
  – client tier
  – middle tier
  – server tier
SAS Intelligence Platform

Client Tier
- SAS DI Studio
- SAS OLAP Cube Studio
- SAS Management Console
- SAS Information Map Studio

Middle Tier
- SAS Enterprise Guide
- SAS Add-In for Microsoft Office
- SAS Web Report Studio
- SAS Information Delivery Portal

Server Tier
- SAS®9 Foundation
- Metadata Server
- Stored Process Server
- OLAP Server
- SAS/CONNECT Server
- Workspace Server

Web Infrastructure Kit
- HTTP Server
- SDK
- webDAV Server
- Java Servlet Container
- Web Infrastructure Kit

© Copyright 2006 ThotWave Technologies, LLC
SAS 9 Clients

SAS® Management Console
SAS® DI Studio
SAS® Enterprise Guide 3.0
SAS® Enterprise Miner 5
SAS® Web Report Studio
SAS® Information Map Studio
SAS® XML Mapper
SAS® OLAP Cube Studio
SAS® Information Delivery Portal 2.0
SAS® Web Report Studio
But clients alone are just not enough!!

- They need to connect to something
- They need to have a "lookup" location to know where to get information when the user selects access to something.. Transparently!
- So what do they connect to?
- All clients connect to a SAS Application Server!
The server tier is the machine where one or more SAS servers is installed and accessed by the BI tools.

There are different types of SAS servers, including:

<table>
<thead>
<tr>
<th>Server Type</th>
<th>Function Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metadata Server</td>
<td>Enables centralized metadata delivery and management to SAS applications across the enterprise.</td>
</tr>
<tr>
<td>Workspace Server</td>
<td>Executes SAS code on behalf of the client applications.</td>
</tr>
<tr>
<td>Stored Process Server</td>
<td>Executes and delivers results from SAS Stored Processes.</td>
</tr>
<tr>
<td>OLAP Server</td>
<td>Delivers pre-summarized “cubes” of data to OLAP clients.</td>
</tr>
</tbody>
</table>
SAS Drug Development

Source Systems

Clintrial eTrial
OpenText
eClinical
Oracle
CROs
Other

ETL q (Defined Business Rules)

Metadata Management

SDD Platform

Version Control
Analysis and Reporting
Biomedical Trial Data Warehouse
Regulatory Submission and Document Management

21 CFR Part 11 Compliance

Data Mining
Reporting
Ad Hoc Querying
OLAP/Cubes

Web Hosting

Portal

Web Server
Desktop

Electronic Submissions

Data Quality, Cleansing, Validation, Aggregation, Enrichment

Access Engines

© Copyright 2006 ThotWave Technologies, LLC
Common Data Integration Tasks

- **Extraction**
  - Data importing
    - Raw files (e.g., labs)
    - SAS Transport
    - Oracle
    - CDISC (e.g., ODM)

- **Transformation**
  - Filtering
  - Data cleansing
  - Categorization & Enrichment
  - Dictionary coding

- **Loading (derived and analysis datasets)**
  - Summarization (for data management)
  - Exporting Data
    - SAS Transport
    - ODM
    - SDTM
Common “BI” Tasks

• Clear separation of data management and reporting (e.g., Reporting is “one PROC away”)
• Derived (or analysis) datasets prepared for us by upstream processes (e.g. DI Studio)
• Delivery or consumption of information:
  – Medical writers
  – Data management
  – Biostatisticians
  – Medical professionals
Which way to go with SAS...

SAS Drug Development®
Enterprise Strategy

Corporate Business Intelligence and Analytics

Data Quality Tier
Ensure data collection and aggregation conforms to master data rules

Data Interface Tier
Ensure consistent mechanisms for interfacing with information sources

Regulated Clinical Business Intelligence and Analytics

*S Image provided by SAS
Using BI for Clinical Research

Overview

SAS 9

Examples

Choosing the right environment

Summary
What to use?

• Who the users are?
  – Stat/programmers
  – Medical writers
  – Data management
  – Biostatisticians
  – “other” departments
What is the purpose?

- Information dissemination
- FDA submissions
- Client review of data
- Production versus ad-hoc
• Production state?
  – Production versus ad-hoc
What to use?

- Existing or New Applications?
- Web Publishing
- Level of Interactivity
What to use?

What kind of processing?
- SQL Queries
- SAS programs
- Macros
• Available Skill Sets?
  – HTML
  – SQL
  – SAS programming
  – Java
  – Excel
  – Web browser
Information Based Medicine will require unprecedented access to diverse, integrated information

1. Patient Information

- Hospital events: admission, surgery, recovery, discharge
- X-rays, MRI, mamograms, etc
- Clinical Record
- Expression Arrays (various tissues)
- Personal genomics
- Analysis lab notes

Challenges
- Volume and complexity of data
- Integrating massive volumes of disparate data
- Need for sophisticated analytics
- Growing collaboration across ecosystem
• Understanding the value of the “platform” approach
  - Metadata
  - Appropriate access to data
  - Lineage & Impact analysis (for DI)
  - Versioning (for some products)
  - Speed time to market
    • Metadata use and reuse
    • Multi-developer capabilities
    • Template reuse (add-in for EG, stored processes)
    • Validation support (e.g., independent programming)
    • Quick starts for unfamiliar tasks
How to reach us...

Greg Nelson
CEO and Founder
greg@thotwave.com