
DataFax Evaluated

- Its Strength, Weakness and Future Necessity

Hanming Tu
Octagon Research Solutions, Inc.
585 East Swedesford Road, Suite 200
Wayne, PA 19087
Web: <http://www.octagonresearch.com/>



Agenda

- Issues
- CDMS Selection
- CDMS Evaluation
- DataFax Strength
- DataFax Necessity
- Conclusion
- Extra: CDISC Roadmap

The Issue

- Octagon had a web-based CDMS with one limitation – no front-end forms for making post-production changes
 - Web-based user interface is friendly
 - Backend relational database SQL server is adequate
 - But post-production changes are inevitable
 - It is the 2nd clinical data management system at Octagon
- The vendor would make post-production changes for us with charges: it is costly and untimely

Temporary Solution

- My first job in Octagon was to develop six stored procedures to make some post-production changes in t-sql:
 - Increase field length
 - Add an item to backend tables
 - Add an item onto front-end forms
 - Add a visit to the study with a few forms
 - Move a form from one visit to another
 - Rename items
- We had to clone the production database, create a test environment, and have our validation team completely test and validate the stored procedures before they are used in the production system

Evaluation: Team

- It was inevitable that we needed a new clinical data management system
- An evaluation team was formed, consisting of members from CDM, Validation and IT departments
- The search began in May 2006 and ended in November 2006

Evaluation: Process

- Send out the Request For Information to about two dozen vendors
 - Contains 108 questions
 - Has the following sections: overall company, EDC access, Paper-based process, application architecture, validation/regulatory, security, study design, dictionary coding, training, reports, audit trail, external data, data transfer/exports, archiving, support, backup and recovery, miscellaneous, and demonstration
- Score each of them based on our business requirements and criteria
- Select finalists based on the functionality, risk, compliance and cost analyses
- Conduct detailed comparison analysis on the finalists

Evaluation Criteria

- **Operational Requirements: total 34**
 - **CRF Design (13):** Interface friendliness, Control Re-usability, Change Traceability, Reporting Capability
 - **Edit Checks (9):** Abilities to create edit checks across CRFs, to create multiple field dependant edit checks, to modify edit checks on the fly without taking study or CRF down.
 - **Permissions and Study Administration (5):** Abilities to manage user and site security and to track setup, structure and data changes.
 - **Query Generation (2):** Abilities to list the history of all queries for each field by date, time, user and query text in color coded and textual query designations
 - **Reports (5):** Abilities to generate a comprehensive edit check report spec of all edit checks across all CRFs and to create canned, ad hoc custom reports.
- **Regulatory Validation: total 13**
 - **SOP (13):** To assess how the software was developed per plan, methodology, and SOPs; how the software to handle data integrity, record retention, patient confidentiality, error logs; and whether it meets the 10/16/2003 HIPAA regulation for data submission and 21 CFR part 11 compliant

Evaluation Criteria

- **Training: total 13**
 - **End User Training (11):** To assess whether the system provides manuals, online help and functional help and whether the vendor can provide adequate onsite or remote trainings to investigator, site coordinator, and monitors.
 - **Technical Training (2):** To assess whether technical training available in one of the following formats: Instructor off-site location; online Training; in-a-box CD-ROM (self paced) and whether the vendor can provide comprehensive manuals.
- **Risk Analysis: total 26**
 - **Financial Stability (7):** To assess the vendor's growth, balance sheet, partnerships, etc.
 - **Management Team (6):** To assess corporate structure, experience, years in Business, technical Expertise, etc.
 - **Customer Metrics (4):** Number of Customers, references, largest installation, etc.
 - **Competitive Positioning (2) :** No Proprietary Technology, Reputation
 - **Facilities (3):** Number of Locations, Adequacy of Locations, local company
 - **Technical Risk (3):** Interoperability, Ability to Integrate, Redundancy
 - **General (1):** FDA Inspection History if inspected

Evaluation Criteria

- **Cost Analysis**
 - ASP
 - ONE-TIME: Training, Set Up/Implementation, Data Transfer, Hardware, Consulting, Software, etc.
 - ON GOING: Annual Fee, Data Transfer, Page/Field Rate, Support, etc.
 - License
 - ONE-TIME: Training, Data Transfer, etc.
 - ON GOING: Software Maintenance, Support, etc.
 - CAPITAL EXPENSE: Core Software, Additional Software, Set Up/Implementation, Hardware, Consulting, etc

Evaluation: POE

- **Process of Elimination**
 - Received 16 RFI responses
 - Based on the scores, we selected 6 from them to have a closer look and/or conduct comparison
 - DataFax was added to the finalists after IT and CDSI addressed the concerns of the CDM users.
 - One system was dropped because it is a SAS-based system and it uses similar technology to one of the earlier systems used by the company
 - Only two systems entered into the final comparison

Comparison: Tasks

- Performed 11 scenarios with 35 tasks in a mock study:
 - Screen Design (3)
 - Edit Check Design (3)
 - Data entry (2)
 - Data entry, multiple users (3)
 - Running edit checks (3)
 - Creation and tracking DCFs (3)
 - Post-production changes (6)
 - Derivations (3)
 - SAS extracts (2)
 - Locking database (3)
 - Remote entry (4)
- Each scenario and task were performed, timed and measured with the cost based on the skill levels

Comparison: Performance

User Performance Metrics From Application Reports			
	DataFax	Competitor	Xsession
Over VPN (*1)			
Local (Seconds/Page)	48.87	51.40	
India (Seconds/Page)	55.39	184.58	
Over Citrix (*2)			
Local (Seconds/Page)	N/A	20.11	
India (Seconds/Page)	56.80	75.13	

User Performance Metrics From India Reports			
	DataFax	Competitor	Xsession
Over VPN			
1. Time for launching the program	3 Sec.	5 Sec.	10 Sec.
2. Time for logging into the system	4 Sec.	30 Sec.	14 min.
3. Time for starting a study	10 Sec.	2 min.	1 min.
4. Average time for entering a single page	54 sec. (*4)	70 Sec. (*5)	30 Sec.
5. Average time for entering a single page	54 sec. (*4)	70 Sec. (*5)	30 Sec.
Over Citrix			
1. Time for launching the program	3 Sec.	3 Sec.	2 Sec.
2. Time for logging into the system	2 Sec.	5 Sec.	5 sec.
3. Time for starting a study	5 Sec.	10 Sec.	5 Sec.
4. Average time for entering a single page	54 sec. (*4)	70 Sec. (*5)	30 Sec.
Open	2 sec	2 sec	
Enter	42 sec.	50 Sec.	
Save	1 sec	2 sec	

Comparison: Summary

- DataFax is performing better than the other system
- A few concerns that CDM users had are in the following areas:
 - Database design
 - Double Data Entry
 - Edit checks
 - Discrepancy management
 - Post-production management
- DataFax is a fax-based system with web interface through iDataFax
- It also has a DDE module for paper-based studies.
- We had focused on evaluating DataFax's DDE capabilities.

DataFax Strength

- **Fax-based System**
 - Simple technology: easy to use and omnipresent
 - Maintain both images and data records for a CRF
- **Intelligent Character Recognition (ICR)**
 - Route CRFs automatically
 - Save time for data entry
- **Multiple Ways for Collecting Data**
 - Fax image
 - Email PDF and image files
 - iDataFax: web-based data validation

DataFax Strength

- **Real Time Capability**
 - Data records are entered into the system through ICR while CRFs are faxed into the system
 - Automated fax-back problem reports for overdue visits and missing pages
 - Edit checks can be set up to alert SAE in real time
- **Flexible Data Management**
 - iDataFax provides the flexibility for sites to be involved in managing CRFs.
 - Monitors, sites and CROs can see the same reports through iDataFax

Future: Adopting a Standard

- **Better to use a common standard**
- **Octagon as a company**
 - Embrace the standard in every direction:
 - A CDISC corporate sponsor
 - A CDISC registered solution provider
 - A primary trainer for CDISC study data tabulation model (SDTM)
 - Push the standard adoption up-stream
 - Regulatory operation: ViewPoint - electronic Common Technical Document (eCTD)
 - Data Integration and Standardization (DIS): OWB - SDTM
 - Data JumpStart: ViewPoint Fuse (EDC) – SDTM
 - (DJS) 26 Standard SDTM compliant CRFs to use as a Base to any study build
 - Creates the ability to have “almost compliant” SDTM datasets out of an operational database
- **DataFax as a system**
 - What to adopt?
 - How to adopt?

Future: Recommended Standards

- FDA recommends CDISC SDTM
 - Federal Register / Volume 71, No. 237 / Monday, December 11, 2006
- PDUFA IV IT Plan (Dec 2007)
 - IT Plan: <http://www.fda.gov/OHRMS/DOCKETS/98fr/07d-0481-gdl0001.pdf>
 - Target Enterprise Architecture (TEA)
 - Business architecture
 - Data architecture
 - Applications architecture
 - Security architecture
 - Standards profile
 - Data Standards: CDISC→HL7
 - Exchange – CDISC SDTM
 - Terminology – CDISC CDASH
 - Clinical Data Repository: JANUS data warehouse

Future: Improving Interoperability

- DataFax CDISC SDTM compliant
 - Export plate data into SDTM domain data sets
 - Import SDTM data sets
- Method
 - Design study according SDTM such as using Data JumpStart
 - Enable the interoperability by adding a metadata layer. Enable it through outside modules may not be a good way for DataFax as I presented in 2003 (<http://www.datafax.com/UserGroup/2003/presentations/Tu.pdf>).
- Possible approach- adding metadata layer
 - Introduce lookup files to define mapping and aggregating DataFax plates and variables
 - Redesign DataFax export tools to use the information in the lookup files to export to SDTM domains and variables
 - When importing, the tools reference to the lookup files as well

Future: Client Proposals In 2004

- **SQL Loader**
 - Merge multiple plates into one table
 - Create tables for study configuration info, audit trail records and DFX_schedule
 - Use CDISC approach to partial dates
 - Shadow field for missing value codes, value labels, partial date imputation info.
 - Allow merge operation across multiple studies
- **Schema**
 - Modules
 - Version control
 - CDISC – write ODM and SDM

Future: Enhancing Architecture

- **Application Integration Architecture (AIA)**
 - An open, standards-based platform for business process management across DataFax, third-party and custom applications
- **Method**
 - Add an application middle tier using AIA
(http://www.datafax.com/UserGroup/2004/presentations/Premier_HanmingTu.PDF,
http://www.datafax.com/UserGroup/2004/presentations/Premier_Vandegrift.PDF)
 - Enable all the DataFax reports and tools in web services
(<http://www.datafax.com/UserGroup/2000/presentations/Andrus.pdf>,
<http://www.datafax.com/UserGroup/2000/presentations/Tu.pdf>)
- **Possible approach – adding open API**
 - Create DFsdm tool to export and import DataFax plates in SDTM
 - DFsdm should combine the functions of DFsqlloader, DFsas DFimport.rpc, and DFexport.rpc
 - DFsdm exports and imports SDTM domains in XML, SAS, or CSV.

Future: Client Proposals In 2004

- **Architecture Enhancement**
 - Firm up the supporting structure – updating the server
 - Simplify client access to server
 - Remove the need for network file system (NFS)
 - Eliminates problems with UNIX file system permissions
- **Security Enhancement**
 - Server enforces all database integrity
 - Structural soundness
 - Internal consistency
 - Server provides all access to study files
 - More detailed specifications and role-based user permissions

Future: Improving Flexibility

- **The foundation of flexibility relies on standard-based, modular and open architecture**
- **Method**
 - Publish open APIs
 - Allow users to add modular functions to the system
- **Possible applications**
 - AE/SAE management
 - Supplies management
 - Site and payment management
 - HTML/PDF reporting and publishing
 - Workflow and task management Standard-based

Future: Client Proposals In 2004

- Higher limits
 - Alpha-numeric study IDs (> 255 studies)
 - 9999 centers
 - Allow visit number on a plate to be a date
- QC notes
 - Allow QC at patient, visit, plate and field levels
 - Allow multiple QCs per patient, visit, plate and field
 - Ability to link one QC to multiple data fields
 - Ability to direct QCs to individuals at sites

Conclusion

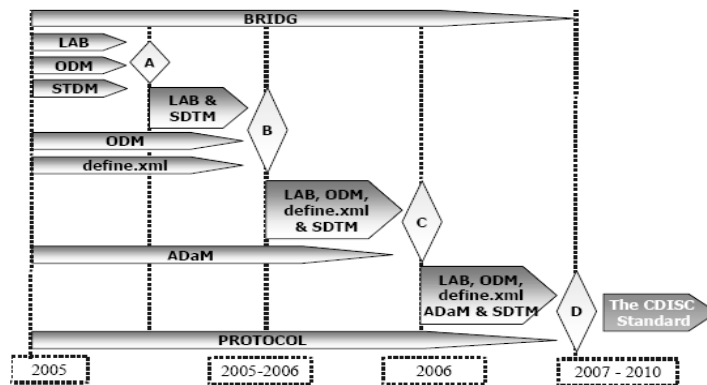
- DataFax is a very stable and solid system with some limitations
- DataFax system can become a viable CDMS by
 - Adopting CDISC standards
 - Adding metadata layer to improve interoperability
 - Utilizing application integration architecture
 - Improving flexibility through standard-based, open and modular approaches
- It is a smart way to gain synergy from user community if a common platform is available

Extra: CDISC Standards and Models

Standard	Description	Implementation Version Release Date
SDTM, SEND	Ready for regulatory submission of CRT 4,000 downloads as of mid-2005	2004*
ODM	CDISC Transport Standard for acquisition, exchange, submission (define.xml) and archive	2001*
LAB	Content standard – available for transfer of clinical lab data to sponsors	2002
ADaM	General Considerations document and examples of datasets for submission	2004
Protocol Representation	Collaborative effort to develop machine-readable standard protocol with data layer	In progress
Terminology Codelists	Developing standard terminology to support all CDISC standards	2006 (Pkg1)
Define.xml	Case Report Tabulation Data Definition Specification	2005*

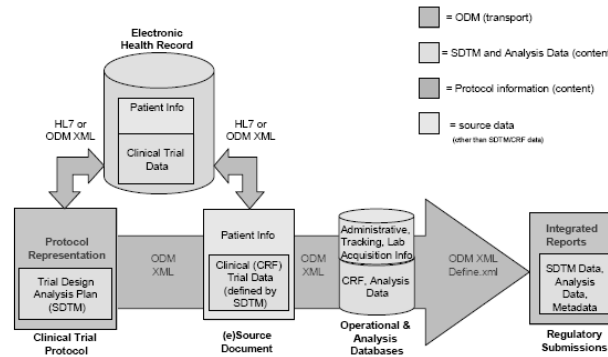
* Specification in FDA Guidance
 Source: CDISC Activities and CDISC Initiation of the CDASH Project, CDASH Webinar 22 Jan 2007, Rebecca Kush
 SEND - Standard for Exchange of Non-clinical Data
 ADaM - Analysis Dataset Model

Extra: CDISC Roadmap Timetable & Milestones



Source: CDISC Roadmap Discussion Document, January 2006.
 Note: BRIDG - Biomedical Research Integrated Domain Group
 ODM - Operational Data Model

Extra: CDISC Data Flow



Source: CDISC Roadmap Discussion Document, January 2006.

Extra: Study Data Tabulation Model

- The SDTM provides a general framework for describing the organization of information collected during human and animal studies.
- The model is built around the concept of observations, which consist of discrete pieces of information collected during a study. Observations normally correspond to rows in a dataset.
- Each observation can be described by a series of named variables. Each variable, which normally corresponds to a column in a dataset, can be classified according to its Role.
- Observations are reported in a series of domains, usually corresponding to data that were collected together. A domain is defined as a collection of observations with a topic-specific commonality about a subject.

Extra: SDTM History

- From SDDM to SDTM
 - Oct 2000: Submission Data Domain Model (SDDM) v1.0
 - Jun 2001: Submission Data Domain Model v1.1
 - Nov 2001: Submission Data Domain Model v2.0
 - Apr 2003: Clinical Trial Data Regulatory Submission Model through Health Level 7
 - Jun 2003: Submission Data Domain Model v3.0; Participation in FDA pilot for patient profile viewer.
 - Jun 2004: SDDM v3.0 → SDTM v1.0
 - Jul 2004: SDTM Implementation Guide v3.1
 - Jan 2005: SDTM draft v1.1
 - Apr 2005: SDTM final v1.1
 - Aug 2005 SDTM IG v3.1.1

Extra: What SDTM Offers

	Before SDTM	After SDTM
Domains	Yes	Yes
Standard Domain Names	No	Yes
Standard Structure	No	Yes
Standard Variables	No	Yes
Standard Variable Names	No	Yes
Standard Terms	No	not yet, but coming (CDASH)

Source: The Review of SDTM Datasets at CDER: A Clinical Reviewer's Perspective by Armando Oliva, MD, Associate Director for Policy, Office of New Drugs, Center for Drug Evaluation and Research, Food and Drug Administration

CDASH - Clinical Data Acquisition Standards Harmonization – content standard
<http://www.cdisc.org/standards/cdash/index.html>

Extra: SDTM Benefits

- No longer having to submit separate patient profiles or listings with a product marketing application.
- Reviewers can now be trained in the principles of standardized datasets and the use of standard software tools, and be able to work with the data more effectively with less preparation time.
 - Standard domain names makes it easy to find the data
 - Standard structure makes it easy to understand the data
 - Standard variables and variable names allows reviewers to be immediately familiar with the data
 - Develop standardized review strategies and tools to store, display, analyze the data
 - Minimal learning curve and consistence
 - Time efficient
- It supports the FDA's efforts to develop a repository for all submitted studies and a suite of standard review tools to access, manipulate, and review the study data.

Q and A



Hanming Tu
Email: htu@octagonresearch.com
Octagon Research Solutions, Inc.
Web: <http://www.octagonresearch.com/>