

## Strategies for Implementing SDTM and ADaM Standards

Susan J. Kenny, Maximum Likelihood Solutions, Inc and  
Octagon Research Solutions, Inc., Chapel Hill, NC  
Michael A. Litzsinger, SCHWARZ BIOSCIENCES, Inc., Research Triangle Park, NC

### ABSTRACT

With the 2004 release of CDISC SDTM and ADaM standards, members of the pharmaceutical industry are all asking the same questions. How and when does the creation of SDTM files occur during the clinical trial process? What files should be used to create analysis files? How do I keep track of all of the data transformations? This presentation will offer some strategies and considerations on how to implement CDISC standards within a pharmaceutical organization. Advantages and disadvantages of implementation strategies will be discussed. A discussion of the types of software solutions that can be used to perform the transformation of data into the SDTM is presented. An illustration of the use of Base SAS® to facilitate the creation of a Findings domain from a database management system to the SDTM or ADaM standards while maintaining important metadata is provided. Ideas, not answers, will be shared to help you think about a CDISC implementation plan for your organization

### INTRODUCTION

It is widely recognized that standards improve process efficiency, regardless of the industry. To that end, the Clinical Data Interchange Standards Consortium (CDISC) has been committed to the development of industry standards to support the processing of clinical trials data over the past 8 years. In July 2004, CDISC released the production version of standards for the design and content of clinical trial tabulation datasets submitted to regulatory authorities, such as the US Food and Drug Administration (FDA). These study data tabulation models (SDTM) specifically address standards for the submission of data typically described as the CRF data. These models have been endorsed by the FDA and are gaining acceptance within the pharmaceutical industry. In December 2004, the Analysis Data Model (ADaM) team of CDISC released a guidance document describing the general considerations for the creation, content, and associated documentation for statistical analysis datasets. These are datasets that are specifically designed to facilitate the statistical analysis and production of study results.

For most submission, both the CRF data and the analysis datasets are submitted to the FDA as part of a new drug application. Now that the CDISC standards have been developed and endorsed by the FDA, many companies are re-engineering their internal processes to adopt them. As complex as the development of the standards was, the implementation will prove to be equally as complex. This paper will discuss some of the issues to consider when implementing both the SDTM and ADaM standards. Suggestions presented here should be considered as such since each organization must develop an implementation roadmap that best fits their environment.

### STUDY DATA TABULATION MODEL REVIEW

A brief review of the key characteristics of the Study Data Tabulation Model (SDTM) is provided here. Since a complete understanding of these models is necessary for successful implementation, readers are encouraged to become familiar with the published CDISC documentation describing the fundamentals of the SDTM and its implementation. The SDTM model, the Implementation Guide, and review comments can be found at <http://www.cdisc.org/models/sds/v3.1/index.html>.

The purpose of the SDTM is to guide the organization, structure, and format of the tabulation data that are to be submitted as part of a product application to a regulatory authority. Tabulation datasets describe the essential data collected during a clinical trial and are one of the four types of data currently submitted to the FDA. The other types of data are patient profiles, listings and analysis datasets. The anticipation is that by submitting the tabulation datasets in a standard structure, the regulatory need for patient profiles and listings will be reduced.

The SDTM is built around the organization of the observations collected about subjects who participated in a clinical study. These observations are organized into series of domains. A domain is defined as a collection of observations that share a common topic. Note that the data organized into one domain may have been represented on one or more case report forms and conversely, that data collected on one case report form may be split into more than one domains. It is important to recognize that where data is collected on the CRF pages and how the data is represented in the study report tables are not factors for deciding into which domain the variables will be placed. Instead, variables are placed into domains according to their topic.

There are three general domain classes where the majority of observations collected during a study are described. These classes are Interventions, Events, and Findings. An Intervention class contains observations relating to treatments or procedures that are intentionally administered. The Exposure domain is an example of an Intervention class. An Events class contains observations relating to occurrences or incidents that happened during the study. The Adverse Events domain is an example of an Events class. A Findings class captures observations resulting from planned evaluations conducted during the study. The Laboratory domain is an example of a Findings class.

In addition to these general domain classes, several special purpose domains are specified in the SDTM. The Demographics domain describes the essential characteristics of the study subjects such as treatment assignment and study start and stop dates. The Comments domain describes a fixed structure for recording free-text comments. The Supplemental Qualifier domains play an important role in capturing variables that cannot be mapped into the standard domains.

The Submission Data Standards (SDS) team of CDISC has detailed the structure and content of over 20 typical domains. These domain models are detailed blueprints of how the data should be represented, the variables to include in the domain and their attributes. Of importance is the assignment of key variables, which uniquely describe an observation.

Successful implementation of the SDTM implies conformance to the defined standards. Conformance is important because it provides the cornerstone for the development of a well-defined data warehouse. With the creation of a data warehouse, the review of new data by both regulatory agencies and sponsors will be facilitated. As detailed in the SDTM Implementation Guide, conformance with the SDTM domain models is indicated by:

- Following the complete metadata structure for data domains and variables. This implies that no additional variables can be added to the model.
- Following the CDISC domain models were applicable
- Including all required and expected variables defined by CDISC
- Using the CDISC specified domain names and prefixes, standard variable names, standard variable labels, and data types for all variables
- Following CDISC specified controlled terminology and format guidelines for variables, when provided
- Ensuring that each record in a dataset includes a set of keys and a topic variable.

## **ANALYSIS DATA MODEL REVIEW**

As part of the process of completing a clinical trial report, analysis datasets (AD) are typically developed from the collected clinical trial data and used to create statistical summaries of efficacy and safety data. These AD's are characterized by the creation of derived analysis variables and/or records. These derived data may represent a statistical calculation of an important outcome measure, such as change from baseline, or may represent the last observation for a subject while under therapy. Depending on the nature of the analyses, these derivations can be complex and use a series of analysis decisions applied to the clinical trial data. These are the analysis decisions that are detailed in the study protocol and/or the statistical analysis plan. The creation of AD's facilitates the statistical analysis and table summaries of study results. As such, these datasets are one of the types of data sent to the regulatory agency to aid in their review.

The CDISC Analysis Data Model (ADaM) defines a standard for AD's to be submitted to the regulatory agency. The underlying principle of these models is to provide clear and unambiguous communication of the content, source, and quality of the datasets submitted in support of the statistical analysis performed by the sponsor. The model concepts are summarized in a General Considerations document and can be found at <http://www.cdisc.org/models/adam/V1.0/index.html>. In most implementations, the AD's contain variables from multiple SDTM domains and contain derived variables for specific analyses. In ADaM, the descriptions of the AD's build on the nomenclature of the SDTM with the addition of attributes, variables and data structures needed for statistical analyses. To achieve the principle of clear and unambiguous communication relies on clear AD documentation. This documentation provides the link between the general description of the analysis found in the protocol or statistical analysis plan and the source data. Of high importance is the clear description of the source(s) of data used as input to the AD's. These descriptions allow the reviewer to trace the derived data items back to their source. Documentation detailing the AD metadata, analysis variable metadata, and the analysis value-level metadata are recognized for their importance. ADaM also defines analysis-level metadata, which describes the major attributes of each important analysis result that is presented in the study report. The purpose of this analysis-level metadata is to allow the reviewer to link from the statistical results to the metadata describing the analysis, the reason for the analysis, and the datasets and programs used to generate the results.

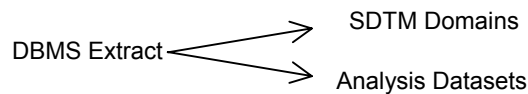
In addition to the principles outlined in the General Considerations document for AD's, ADaM is developing examples of AD's and their associated documentation that relate to frequently used statistical methods. These examples illustrate the application of ADaM principles to a typical analysis.

## THE DEVELOPMENT LIFE CYCLE OF SDTM AND ADAM DATASETS

Before the SDTM standard was developed, the typical scenario for the creation of clinical trial datasets was to create an extract from the database management system (DBMS), such as OracleClinical and prepare this extract as the submission tabulation files and build analysis datasets from this extract. Now that the SDTM standards are part of this development cycle, the salient question is what will become the new typical scenario? Assuming that both tabulation datasets and analysis datasets and associated documentation will be submitted to a regulatory agency, there are at least four options for the development life cycle of these data, each with advantages and disadvantages. These are described below.

### PARALLEL METHOD

This development path is illustrated as:



The advantages of this method are:

- The development of the SDTM and ADaM datasets are independent and can be completed at any time without input from the other.
- The creation of the SDTM can happen at the time of submission so there is no effort wasted if the clinical trial is either unsuccessful or not included in a submission.
- The independence of the SDTM and ADaM datasets allows for parallel project teams to perform the extraction, transformation, and load (ETL) processes. This may be important for outsourced projects.
- This method requires a minimum amount of re-engineering of existing processes within most pharmaceutical companies.

The disadvantages of this method are:

- Documentation for each set of data share no similarity and parallel creation decreases efficiency.
- Derivation of created variables in analysis datasets does not reference variables in the SDTM. This creates a significant disconnect between the two sets of submission data.
- The regulatory agency does not have the original DBMS extract with which to verify or explore derivations performed in the analysis datasets. Similarly, they would not have the DBMS annotated CRF's to understand the original source of the derivations of the analysis variables.
- Any analysis programs submitted to the agency as analysis-level documentation has limited value since the source data is not available.
- Validation is necessary to ensure that similar records or variables in both SDTM and ADaM datasets are identical, such as indication of which record is considered baseline.

### RETROSPECTIVE DEVELOPMENT

This development path is illustrated as:



The advantages to this method are:

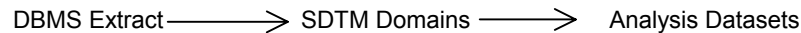
- The creation of the SDTM can happen at the time of submission so there is no effort wasted if the clinical trial is either unsuccessful or not included in a submission.
- As enhancements to the SDTM standards are released, the analysis datasets are not affected and any enhancement can be represented during the creation of the SDTM.

The disadvantages of this method are:

- The regulatory agency does not have the original DBMS extract with which to verify or explore derivations performed in the analysis datasets. Similarly, they would not have the DBMS annotated CRF's to understand the original source of the derivations of the analysis variables.
- Any analysis programs submitted to the agency as analysis-level documentation has limited value since the source data is not available.
- Any date imputation or other types of hard coding performed during the creation of the analysis datasets would have to be undone since the SDTM represent the data as it was collected.
- All CRF variables represented in the SDTM would need to be retained in the analysis datasets even if they are not used for analysis. This increases the complexity of documentation.
- Validation is necessary to ensure that the SDTM adequately represent the original source data. This step could be potentially difficult and result in a loss of efficiency.

## LINEAR METHOD

This development method is illustrated as:



The advantages of this method are:

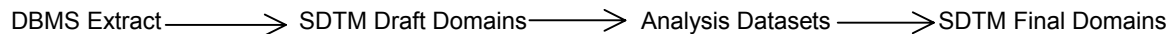
- Analysis programs submitted to the agency as analysis-level documentation utilize the SDTM domains as input and are thus useable and informative to the reviewer.
- Using SDTM domains as input to analysis datasets allows for the standardization of analysis dataset structures and programming methods to produce study report summaries.
- It appeals to the logical flow of the software development life cycle of A to B to C.

The disadvantages of this method are:

- The development of the analysis datasets relies on the completion of the SDTM domains.
- The SDTM domains are created for all clinical trials regardless of whether they will be part of a submission.
- It is potentially more difficult to manage if the data management and/or the biostatistics is outsourced.

## HYBRID METHOD

This development method can be illustrated as:



With this method, the differences between SDTM Draft domains and SDTM Final domains are envisioned to be small. The SDTM Final domains contain the subset of variables or records that are optimally created during the analysis or at the final stage of submission preparation. An example is the creation of USUBJID. This variable is required in the SDTM and provides a unique key identifier for a given subject. In some situations, however, the creation of USUBJID cannot be defined until all studies are complete since a given subject may participate in multiple trials. Other examples include the creation of expected variable that is present in all Findings domains that indicate the data record considered to be the baseline value (e.g. 'EGBLFL'). Since these indicator flags likely would be derived in the AD's, creating the Final SDTM domains retrospectively from the AD's prevents redundant derivation and eliminates the possibility of discord between SDTM domain and the analysis dataset. Finally, population indicator variables, such as those for intent-to-treat or per protocol status, can be optimally created in the AD and then placed in the supplemental qualifier domain.

The advantages to this method are:

- With a few possible exceptions, analysis programs submitted to the agency as analysis-level documentation utilize the SDTM domains as input and are thus useable and informative to the reviewer.
- Using SDTM domains as input to analysis datasets allows for standardization of analysis dataset structures and programming methods to produced study report summaries.
- The variables or records in the SDTM that need biostatistical input, such as indication of baseline records or creation of population flags, is done in harmony with analysis datasets so there is no possibility of discrepancy.
- If important, derived records can be added to the SDTM domains thus providing the reviewer with both CRF and analysis records.
- Final completion of the SDTM domains can be done at the time of submission.

The disadvantages of this method are:

- The development of the analysis datasets relies on the completion of the SDTM domains.
- The SDTM domains are created for all clinical trials regardless of whether they will be part of a submission.
- This may be potentially more difficult to manage if the data management and/or the biostatistics is outsourced.

## RECOMMENDATIONS

Each organization will need to leverage the advantages and disadvantages of these methods when deciding an implementation plan. For submissions that are prepared within the near future, several of the above methods may need to be used in tandem to accommodate both legacy data and ongoing studies. But as CDISC standards become adopted within an organization, one would expect that efficiencies will be gained if one method were used for all new studies going forward. Weighing the advantages and disadvantages of each method above, the linear or the hybrid method are the most parsimonious and are long-term solutions. These methods follow the logical pattern of software

development and they provide the reviewer with unambiguous documentation and source data for the analysis datasets. The parallel and retrospective methods are potentially advantageous as short-term solutions to be used over the years of transition to the linear or hybrid method.

## **SOFTWARE SOLUTIONS FOR THE CREATION OF SDTM DOMAINS**

Upon review of the SDTM, it is apparent that creating the domains directly from an extract of the database management system (DBMS) is unlikely. In addition to restructuring of the data, input from the clinical team will be needed to define expected variables, such as the study reference start and stop dates. Therefore, some amount of post processing of the DBMS extracts will be necessary to create the SDTM domains.

### **VENDOR SOFTWARE**

The reorganization of the DBMS extracts is an example of a typical ETL (Extraction, Transformation, and Load) process currently used in many industries. There exists a variety of marketed software designed specifically to address and automate ETL processes. Certainly, one of the market leaders is the SAS® product ETL Studio (<http://www.sas.com/technologies/dw/etl/etlstudio/>). Using a visual design tool, this product helps organizations build, implement and manage ETL processes from source to destination, regardless of data sources or platforms. In-depth data transformations can be performed with minimal programming. The first commercially available software designed specifically to convert Oracle® Clinical data to the SDTM format is OC2SDS™ from CSSInformatics (<http://www.csscomp.com/web/products/oc2sds.htm>). Additional software, such as WEBSDM™ from Lincoln Technologies (<http://www.lincolntechnologies.com/Technology/standards.html>) is designed to take advantage of the efficiencies of standardized data. By no means is this an exhaustive list of potential software vendors that have products that could be used in the creation of SDTM domains. Organizations are encouraged to search across all products for one that fulfills the necessary functional requirements.

### **BASE SAS PROGRAMMING**

However, if budgets do not allow for the purchase of additional software, the post-processing of DBMS extracts to SDTM domains can be completed using Base SAS software. This is especially feasible if an organization has standard CRF pages that are used across all of their clinical trials. Standard CRF pages form the foundation upon which standard DBMS structure, standard extracts, and standard mapping of the DBMS extract to the SDTM domains can be built. Without standard CRF pages, the efficiency of using Base SAS software for the ETL is reduced yet still practical. Using Base SAS software provides advantages by allowing complete control over the mapping of the domains as well as results in familiar documentation in the form of SAS programs and log files. The SAS code needed for any derived variables or observations that are placed in the SDTM is the same code that would be used in the creation of the AD's. This results not only in programming efficiency but it also streamlines the documentation and increases the harmonization between SDTM domains and AD's.

### **DEVELOPING USEFUL BASE SAS MACROS**

If Base SAS is used to create the SDTM domains, efficiency will be gained by developing a suite of macros to perform many of the redundant transformations that are used across domains. Macros designed with the following functionality are suggested:

- To create core variables used across domains and submissions (key subject identifiers, study reference dates). This could be paired with the automation of the creation of a text document to use as reference for important decisions and definitions specific to these subject or study identifiers.
- To order variables in the SDTM according to the models.
- To obtain variable labels
- To check model conformance with published SDTM domains using an Excel derived SAS datasets. This would include report generation used for validation and cross check of model development.
- To import of trial design dataset Trial Visit (TV) for standard population for domain visit variables (protocol description text, visit number and planned visit day).
- To create the trial design dataset Subject Visits (SV).
- To populate the DS domain with the subject specific protocol and trial milestones obtained from standard variable input from the DBMS extracts.
- Creation of ISO8601 compliant character date strings and durations.
- Calculation of study day.

Readers are referred to other authors who have presented useful SDTM macro code in greater detail (Shostak, 2005)

## AN EXAMPLE OF USING SAS TO CONVERT A DBMS EXTRACT TO A FINDINGS DOMAIN TO AN ANALYSIS DATASET

Two macros will be detailed here because they illustrate an important relationship between SDTM domains and AD's. First, consider the circumstance where a DBMS extract is normalized to adhere to the SDTM table structure. Below is an abbreviated representation of a DBMS extract from the data collected on a typical vital sign CRF (VITSIGN) page.

PATNO	VISIT	HT_CM	WT_KG	BMI	FRMSIZE
015200	END OF STUDY	172.7	54.432	18.246	MEDIUM

Note that the format for height in centimeters is 5.1, for weight in kilograms is 6.3 and the calculation of BMI is formatted as 6.3. Assume that the character variable FRMSIZE was defined with a length of \$6 to accommodate possible values of 'SMALL', 'MEDIUM', 'LARGE', 'UNK'.

When this DBMS extract of VITSIGN is converted to structure of the Vital Sign (VS) Findings domain, these data are normalized and result in following abbreviated representation:

USUBJID	VISIT	VSTESTCD	VSTEST	VSSTRESN	VSSTRESC
100-015200	END OF STUDY	HEIGHT	Height in cm	172.700	172.7
100-015200	END OF STUDY	WEIGHT	Weight in kg	54.432	54.432
100-015200	END OF STUDY	BMI	BMI	18.246	18.246
100-015200	END OF STUDY	FRAME	Frame Size		MEDIUM

Note that the numeric variable VSSTRESN houses all numeric variables and therefore will need to have a numeric format that will accommodate all levels of precision that were used in any of the vital sign test measurements. Since three decimal places were the maximum number used in the original DBMS file, then all values of VSSTRESN will have 3 decimal places. When using SAS, 0's are used to pad any value that does not have this number of decimal values, as illustrated above with the value of 172.700 for Height.

The normalized structure of the SDTM Findings domain is conducive to data warehousing and transmission of data. However, it is not often an ideal structure for AD's since many types of analysis and table summaries are easier to perform if each visit data is on one record. Therefore, to create a by-visit vital sign analysis dataset, the VS domain will be flipped back to a horizontal structure. The structure of an analysis file for change from baseline of vital signs would be as follows:

USUBJID	VISIT	HT_CM	WT_KG	BMI	FRMSIZE	CFB_WT
100-015200	END OF STUDY	172.7	54.432	18.246	MEDIUM	-3.548

It is evident from this example, that a successful 'flipping' of the SDTM domain to the AD requires the knowledge of the original format of the both the numeric and character variables. Without this knowledge, SAS will represent all numeric values with 3 decimal places and, in the absence of 'defensive' programming, will assign the length of the FRMSIZE variable to the length of the value that it first encounters in the database.

Two SAS macros can help manage this process. The first is a macro that creates a list of macro variables that correspond to the DBMS variable names that will become rows in the SDTM domain. This approach to creating an array of macro variables has been presented by other authors (for example, Fehd, 2004). The creation of this 'virtual array' of macro variables is presented here in a simple form:

```
%Macro V_array (Name=Var,
                List=,
                Delimitr = %str( )<%str( )]&!$*%str()^-/%>\ / );
%If "&Name." ne "" and "&List." ne "" and %length(&Name.) le 28
%then %do;
    %Let I = 1;
    %Let Item = %Qscan(&List,&I,&Delimitr);
    %Do %until(&Item. = );
        %global &Name.&I;
        %Let &Name.&I = &Item.;
        %Let I = %eval(&I + 1);
        %Let Item = %Qscan(&List,&I,&Delimitr);
    %end;
```

```

    %end;
    %global Dimarray;
    %Let Dimarray = %eval(&I - 1);
%end;
%Mend;

```

Note that the output from this macro is a list of macro variables, as well as a global variable to indicate the number of macro variables created. The macro variable 'List' specifies the variable names in the DBMS extract that will eventually be transposed. Thus, for the above example, the macro call would be:

```
%v_array(list=ht_cm, wt_kg, bmi, frmsize);
```

This results in the creation of four macro variables, &var1 - - &var4, whose values are the variable names and &dimarray with a value of 4. Once these are created, they can be used as input to other macros, such as the following macro, which saves the value-level metadata of these variables in a SAS dataset.

```

%macro savemeta(libname=&oc, dsn= , list=, metafile=) ;
** obtain value level metadata for variables that will be flipped **;
proc sql;
  create table &metafile as
  select name, type, label, format
    from sashelp.vcolumn
  where libname=%upcase("&libname") and memname=%upcase("&dsn")
    and name in (%do _i=1 %to &dimarray;
                  %upcase("&&var&_i")
                %end;
                );
quit;
%mend;

```

Following the example above, this macro call would be:

```
%savemeta(dsn=VITSIGN, list=&list, metafile=vs_meta)
```

The SAS dataset 'VS\_META' would contain four records with values of the original variable name, type, label, and format that were the metadata specified in the DBMS data table. This data table would be used as input to the documentation for the creation of the analysis dataset and most importantly can be used to automate the creation of SAS LENGTH and FORMAT statements used in the AD creation program. Saving this metadata in a SAS table ensures that the proper value-level metadata is used when the AD is created without user input.

Using the macro variables created in %v\_array can be used to macroitize the transformation of the VITSIGN file to the normalized SDTM VS domain. An excerpt of example code is as follows:

```

%do _i=1 %to &dimarray;
  &prefix.testcd=vname(&&var&_i);
  &prefix.test=vlabel(&&var&_i);

  *parse text between [ ] to obtain units *;
  _length=index(vlabel(&&var&_i), '[')-index(vlabel(&&var&_i), '[')-1;
  if _length>0 then
&prefix.orresu=substr(vlabel(&&var&_i), index(vlabel(&&var&_i), '[')+1, _length);

    if vtype(&&var&_i)='C' then do;
      &prefix.stresc=left(&&var&_i);
      &prefix.orres=left(&&var&_i);
    end;
  else if vtype(&&var&_i)='N' then do;
    if &&var&_i>.z then do;
      &prefix.stresc=left(&&var&_i);
      &prefix.orres=left(&&var&_i);
      &prefix.stresn=input(&prefix.stresc,8.);
    end;
  end;
  if &prefix.stresc>' ' then output;
%end;

```

This code excerpt uses the metadata (VNAME and VLABEL) of each of the variables to be transposed and creates both the character and the numeric representation of any SDTM Findings domain variables. In this example, the label of variable in the DBMS VITSIGN file is standardized and contains the value of the original unit of the measure within brackets (for example "Height [cm]"). Parsing the label yields the SDTM Findings variable 'xxORRESU'. This is an illustration of how simple standardization of the DBMS metadata can result in efficiency gains during the creation of the SDTM domains.

## CONCLUSION

It is important to recognize that the implementation of standards, in any industry, will bring about change. This change not only will be manifested within organizations but also within the standards themselves. As the SDTM is used by more organizations, enhancements to the models will be made. Other CDISC standards, such as the specification of the DEFINE.XML schema, will influence both the SDTM and AD. Using XML as the document format for the data definition tables will result in machine-readable metadata for the table of contents of data and the accompanying data description tables included in a submission. The power of XML has the promise to revolutionize the management of clinical trial information. For example, the illustration above of saving the value level metadata for later use in the creation of the analysis datasets will become obsolete once XML is standard since this information will be in the DEFINE.XML and can be machine read during any ETL process.

The FDA has endorsed the CDISC standards (<http://www.fda.gov/bbs/topics/news/2004/NEW01095.html> ) and is anxious for the industry to begin to apply the standards to submission data. Waiting for these standards to never change or waiting for answers on the best way to proceed will be an exercise in futility. When developing a process for the implementation of SDTM and ADaM standards, a willingness to allow creativity and trial and error to be part of this process will be of benefit since there is neither yellow brick road nor well-trodden path to follow. On face value it may appear that implementation of any standard is restrictive but upon thorough examination of the SDTM and ADaM, it is clear that the models provide a great deal of flexibility. Perhaps the words of the Scott Adams, of Dilbert fame, best summarize the strategy for implementation: "Creativity is allowing yourself to make mistakes. Art is knowing which ones to keep". Indeed, implementation of CDISC standards is both an artistic science and a scientific art.

## REFERENCES

Adams, Scott. 1997 "The Dilbert Principle". HarperBusiness.  
Fehd, Ron . 2004 "Array: Construction and Usage of Arrays of Macro Variables". SUGI29 Proceedings.  
Shostak, Jack. 2005 "Implementation of the CDISC SDTM at the Duke Clinical Research Institute". PharmaSUG Proceedings.

## RECOMMENDED READING

Visit [www.cdisc.org](http://www.cdisc.org) to read:

The Study Data Tabulation Model (SDTM), June 25, 2004  
The Study Data Tabulation Model Implementation Guide (SDTM-IG), July 14, 2004  
Statistical Analysis Dataset Model: General Considerations Version 1.0, December 2004  
Case Report Tabulation Data Description Specification (define.xml), February 2005  
Review and participate in the Discussion Forums (<http://www.cdisc.org/discussions/discussions.html> )

## CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Susan J. Kenny, PhD  
Maximum Likelihood Solutions, Inc.  
PO Box 2074  
Chapel Hill, NC 27515  
Email: [susankenny@mebtel.net](mailto:susankenny@mebtel.net)

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.