

Defining the Development Process and Governance of Implementing ADaM within an Organization

Chris Decker, d-Wise Technologies, Cary, NC, USA

ABSTRACT

Over five years ago, the CDISC ADaM standard was released as a set of principles and best practices. In reality, the model was not really a 'standard' but a general set of guidelines for implementing analysis data; therefore companies avoiding implementing the standard. With the release of the ADaM Implementation Guide at the end of 2009 and the FDA's renewed commitment to standards, companies have realized they have to jump on the ADaM bandwagon or get left behind. This paper presents a case study describing an ADaM development methodology, ADaM implementation lessons learned, and an overall (and ongoing) standards governance process. This case study uses experiences across projects with a various companies to combine lessons learned, best practices, things 'not to do' and potential future state to support the development and use of analysis data standards.

INTRODUCTION

Over five years ago, the CDISC ADaM standard was released as a set of principles and best practices. However, in practice, the model was not really a 'standard' but a general set guidelines for implementing analysis data. Given the state of the standard and the lack of any guidance from the FDA, most companies avoided the need to implement ADaM – "if it's not required, I'm not doing it". At the end of 2009, the CDISC ADaM team released the ADaM Implementation Guide version 1.0 containing specific standard data structures, required variable names, naming conventions, rules, and examples of how to use the model. While still including a lot of flexibility, the industry now had the foundation of a standard to implement. With this new guide and the FDA's renewed commitment to standards, companies have realized they have to rapidly begin implementing ADaM or risk behind the curve. This paper presents a case study describing the core components of implementing ADaM within an organization.

- ADaM development methodology – using the analysis defined in the SAP/Protocol to drive ADaM development
- ADaM implementation lessons learned – describe decisions, best practices, and lessons learned of implementing ADaM
- Standards governance – high level overview of a potential standards governance process

This case study uses experiences across projects with a various companies to combine lessons learned, best practices, things 'not to do' and potential future state to support the development and use of analysis data standards.

ITERATIVE ADAM DEVELOPMENT

Beginning the development process for ADaM across your organization can be an overwhelming task. This is even more evident if you try to approach this challenge across the entire organization all at once. The key is to start small and build iteratively defining different levels of standards and refining the standards and process along the way. Just remember that Rome wasn't built in a day and neither will your ADaM standards be built in a day, week, month, or forever. Standards development must be a continuously iterative process that is refined along the way.

The key to defining analysis data standards is to stop focusing on the rows and column data structure and instead concentrate on understanding the analyses you need to produce the results required by the clinical study. The goal isn't to create as many tables, figures and listings as possible but to create the analysis data that can support the critical study endpoints.

Figure 1 outlines the process we have begun to implement with a number of customers that moves towards a focus on the analysis results driving the design and development of the analysis data.

PhUSE 2011

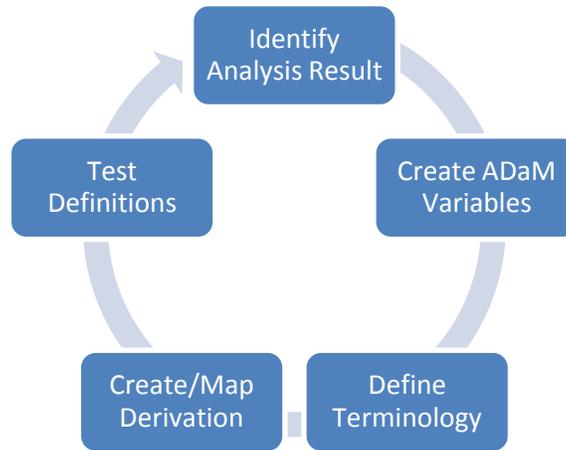


Figure 1. Analysis Package Methodology

IDENTIFY ANALYSIS RESULT

The first step in this process is to define the analysis result of interest. In the current world, we would just start building analysis data sets based on a set of summary table templates that have developed. In a future state, we need to shift to a world where we focus on reviewing the protocol and/or statistical analysis plan and identify specific analysis results that are driven by the list of efficacy and safety reporting requirements and not worry so much about how they fit on a rectangular piece of paper.

In both the current and potential future state, you need to identify and describe very granular analysis result. The demographics table below is a traditional table familiar to all of us.

	Drug A (N=)	Drug B (N=)	All (N=)	
Sex	n (%)	n (%)	n (%)	Analysis Result 1
Male	xx (x.x)	xx (x.x)	xx (x.x)	AR1
Female	xx (x.x)	xx (x.x)	xx (x.x)	
Race	n (%)	n (%)	n (%)	Analysis Result 2
Caucasian	xx (x.x)	xx (x.x)	xx (x.x)	AR2
Hispanic	xx (x.x)	xx (x.x)	xx (x.x)	
Asian	xx (x.x)	xx (x.x)	xx (x.x)	
African	xx (x.x)	xx (x.x)	xx (x.x)	
Age				Analysis Result 3
n				AR3
Mean				
Standard Deviation				
Median				
Minimum				
Maximum				

Table 1. Example of Traditional Summary Table

In our current world, we always think about developing this entire demographics table. Instead, let's think about each of the individual analysis results identified on this table. In this example, there are three analysis results defined, each with different requirements for both reporting and the ADaM data sets. The metadata listed in the table below is just a small sample of what could be captured to describe each analysis result.

Analysis Result Identifier	Analysis Result Label	Reporting Requirements		Data Requirements		
		Type	Stats	Analysis Value	Treatment	Population
AR1	Summary of Sex	FREQ	%	Sex	Planned Treatment	Intent to Treat
AR2	Summary of Race	FREQ	%	Race	Planned Treatment	Intent to Treat

PhUSE 2011

Analysis Result Identifier	Analysis Result Label	Reporting Requirements		Data Requirements		
		CONT	N. Mean, etc.	Age	Planned Treatment	Intent to Treat
AR3	Summary of Age	CONT	N. Mean, etc.	Age	Planned Treatment	Intent to Treat

Table 2. Example of Analysis Result Metadata

While this is a simple example, you can see how each analysis result is separate on the the table and metadata captured to support both the reporting and data needs. The example below is a little bit more in depth and shows additional requirements. In this analysis, a pain score is collected daily, weekly average score calculated, and a simple t-test is used to compare the two treatments

	Drug A (N=)	Drug B (N=)	P-value (N=)	Analysis Result
Weekly Mean Pain Score by Week				
Week 1			P-value	AR1
n				AR2
Mean				
Standard Deviation				
Week 2			P-value	AR1
n				AR2
Mean				
Standard Deviation				

Table 3. Example of Efficacy Summary Table

In this example, we have two basic analysis results. The first is the calculation of the statistical test, and the second is the summary statistics for each week. The table below outlines the information to be collected for each analysis result.

Analysis Result Identifier	Analysis Result Label	Reporting Requirements		Data Requirements					
		Type	Stats	Analysis Value	Treatment	Population	Parameter	Time	Covariate 1
AR4	Comparison of Treatments for Weekly Mean Pain Score	PVALUE	T-TEST	Dose	Planned Treatment	Per Protocol	Weekly Mean Dose	Week	Pain at Baseline
AR5	Weekly Mean Dose	CONT	N, Mean, etc.	Dose	Planned Treatment	Per Protocol	Weekly Mean Dose	Week	

Table 4. Example of Efficacy Analysis Results Metadata

In this example, the analysis result metadata is more complex due to the need for a statistical test and calculation of mean over weeks. As you can see the reporting metadata is different for each analysis result as well as the addition of covariate metadata for the statistical test.

CREATE ADAM VARIABLES

Now that the specific components required to generate an analysis result have been identified, the next step is to define the specific ADaM elements that are needed. Examples of specific ADaM elements required to support the analysis results in the table above are listed the table below.

AR Identifier	Element Type	ADaM Structure	ADaM Variable	Variable attribute 1	Variable attribute 2	Terminology	Algorithm
AR1	Analysis Value	ADSL	RACE	Text	\$20	RACE	DM.RACE
AR1	Treatment	ADSL	TRT01P	Text		TRT01P	Description

PhUSE 2011

AR Identifier	Element Type	ADaM Structure	ADaM Variable	Variable attribute 1	Variable attribute 2	Terminology	Algorithm
AR1	Population	ADSL	ITTFLL	Text	\$1	YN	Description
AR5	Analysis Value	ADPN	AVAL	Float	8.		Description
AR5	Treatment	ADSL	TRT01P	Text		TRT01P	Description
AR5	Population	ADSL	PPROTFL	Text	\$1	YN	Description
AR5	Time	ADPN	AVISIT	Text	\$10	AVISIT	Description
AR5	Parameter	ADPN	PARAMCD	Text	\$8	WKPNSCR	Description

Table 5. Example of ADaM Elements for Efficacy Analysis

Within the actual design and implementation, the table above was divided into multiple metadata libraries, each capturing extensive metadata about the ADaM structures, variables, terminology, and derivations. Once a robust metadata library containing these elements is in place, users would be able to pull from the library to map to the required components. Most of our ongoing projects are still in the process of populating the library across their therapeutic areas. In the case of ADaM structures and variables, a library of controlled terminology and associated process was put in place to control naming and meaning of these elements at different levels within the standards hierarchy.

DEFINE TERMINOLOGY

When controlled terminology is often referenced, only the specific values of ADaM variables are considered. A key component of developing the ADaM elements in the table above is to define controlled terminology not just for the variables values but for all other elements including:

- Analysis result identifiers
- ADaM structures
- ADaM variable names and attributes
- Variable value (typical terminology)
- Value level terminology
- Derivation terminology where possible

Some of the attributes define specific for variable values and value level information included a unique identifier, a unique name, source of the term (i.e. CDISC, internal, other external source), whether the term was extensive, each value within the term, optional aliases for each value, and versioning for each individual value.

CREATE VALUE LEVEL METADATA

Due to the two dimensional nature of ADaM as well as other CDISC models, there is a need to support value level metadata. Value level metadata implies defining the attributes of a value within one variable based on the value of another variable (e.g. attributes of AVAL when PARAM="MEAN PAIN SCORE"). This is often referenced as value level metadata in SDTM or parameter level metadata in ADaM. A parameter level library captures attributes such as:

- Unique parameter identifier
- Parameter name (i.e. PARAMCD)
- Parameter value (i.e. PARAM)
- Linked Variable or the variable to describe based on the parameter value
- Linked metadata attributes associated with the linked variable (format, length, derivation)

The table below is just an example of some information collected within value level metadata.

Parameter Identifier	Parameter Name	Parameter Value	Parameter Code	Linked Variable	Linked Format	Linked Derivation	OTHER LINKED METADATA
PVK1	PARAMCD	Systolic Blood Pressure (mmHg)	ASYSTBP	AVAL	4.	LDERV1	Contains other metadata describing AVAL

Table 6. Example of ADaM Value Level Metadata

The table above captures the parameter value, associated code and the linked variable being described and its associated metadata. In this example, the ADaM element AVAL is being described when the PARAMCD element is ASYSTBP. It is very important that a company manages the controlled terminology around the ADaM elements, especially related to the PARAM/PARAMCD definitions.

CREATE AND MAP DERIVATIONS

PhUSE 2011

One of the most challenging tasks of developing the components that make up an analysis result is defining the derivation associated with the variable or value level definition. Derivations have always been a black box and the ADaM Implementation Guide basically communicates that anything (code, text, links, etc.) can be placed in a derivation as long as it is 'clear and unambiguous'. This provides the greatest flexibility for end users to communicate but unfortunately makes controlling the content nearly impossible. While in the very early stages, a derivation library was designed attempting to capture metadata that would be somewhat usable and potentially machine readable. Below is a sample of metadata captured for derivations.

Derivation Identifier	Derivation Name	Category	Description	Executable Code	Derivation Link 1
DV1	RACE	COPY	Copy from SDTM	DM.RACE	Link to source information
DV2	AGEGR1	DIRECT	Age Grouping <65, >-65	If ADSL.AGE<65 then AGEGR1='AGE<65' Else if ADSL.AGE>=65 then AGEGR1="AGE>=65'	Link to source information
DV3	WEEKMN DOSE	INTERNAL	Weekly Mean Dose calculated over 7 days	%WEEKLYMNDOSE(var1=, var2=, var3=,...)	Link to source information
DV4	TIMETO EVENT	EXTERNAL		%TIMETOEVENT()	Link to source information

Table 6. Example of ADaM Derivation Metadata

In the table above, metadata was captured in an attempt to fully qualify the derivation. The team used category to capture different levels of complexity in defining the derivation and a subset of category values are listed above.

- COPY- derivation is a direct copy and the executable code can be used to copy the variable directly from the source.
- DIRECT – derivation can be applied within a single record on a source data set and the fully executable code is captured.
- INTERNAL – derivation includes values/variables across records but is internal to the data source
- EXTERNAL – derivation includes values/variables across records and data sources

In the case of the INTERNAL and EXTERNAL categories, separate functions (i.e. macros) are being built and validated to support those derivations. The library can contain one or more derivation links for each derivation which includes a machine readable link to an external document, SAS code, or other piece of information to provide traceability and purpose for the derivation. The project is in the early stages of building and piloting this approach.

BUILDING AN ADAM LIBRARY

The different ADaM elements described within the previous sections make up an analysis 'package'. This package includes the analysis results and all the components and elements which are needed to generate a specific analysis endpoint. The columns referenced in the table as identifiers are used to provide relationships between ADaM elements and other metadata tables including terminology, derivations and potentially value level information.

Once the team has an overall understanding of how and what needs to be collected for an individual analysis result, the next step is to continue iterating through this process building a repository of analysis 'packages'. Once this exercise has been completed for a study, a full list of analysis elements will be available and the overall analysis data sets can be constructed. For example, looking across the analysis packages would determine the variables required within ADSL. This will allow you to collate those ADSL variables into a standard ADSL template that can be used across studies and added to as necessary. After completing this exercise within a study, the process can be piloted across similar studies and therapeutic areas adding and refining your ADaM library as you go.

DEFINING AND FOLLOWING ADAM RULES

In the developing an ADaM model for your organization, there are a number of decisions you have to make, and more importantly, make sure you follow those decisions. The 'openness' and flexibility of ADaM is both a blessing and a curse and you must include as much rigor in the process as your creative statisticians and statistical programmers can handle. This section outlines a few high level decisions to consider when developing your process around ADaM.

DEFINING THE DATA FLOW

PhUSE 2011

The first step is to clearly define and reach consensus on how your organization will support the data flow throughout a study lifecycle. There are a number of discussions and decisions that must be made to support your organization's desired workflow.

- Will you always build ADaM from SDTM?
- How much derived data should really be supported within SDTM versus ADaM?
- Do we build large ADaM data sets differentiated by parameters or do we build a lot of small ADaM data sets focused on very specific analyses?
- Should we create TLFs only from ADaM or a combination of ADaM and SDTM?
- When we integrate data, should we integrate SDTM and build ADaM from an integrated SDTM database or should we integrate ADaM?

These are all questions that must be discussed, a consensus reached and a detailed process development to ensure everyone within the organization understands the data flow. The goal of the data flow should be to lock down the process as much as possible to avoid too much flexibility which will inevitably lead to inconsistency analysis standards and data. Although ADaM is still fairly early in its adoption within the industry, our experience has shown the following data flow and guidelines support a robust process and removes the wiggle room from ADaM development. The next few sections are more opinions based on experience with customers and work with the CDISC teams.

SOURCE OF ADAM

Many individuals within the industry ask the question of what should be the source of ADaM data and some propose building ADaM data from the raw EDC or CDMS data because it's 'easier'. While this is technically possible, it is the wrong approach to ensuring continuity and traceability in your data. ADaM data should always be created from SDTM to support the linear traceability of the data. Analysis data sets and associated metadata must clearly communicate how the analysis data was created and standardizing the source of the analysis data supports this communication. When traceability is successfully implemented, users are able to identify the data that comes directly from SDTM, data that is derived or imputed within the ADaM analysis dataset, and the method used to create derived or imputed data used for analyses. The only exception to the source data for ADaM would include reference data sets that do not have a location within SDTM (i.e. laboratory reference ranges). These can be carried along to support derivations within ADaM.

TO DERIVE OR NOT TO DERIVE

The question of what derived data to include in SDTM versus ADaM has been unanswered and debated for many years. One of the reasons for this challenging question is the different maturity levels of each model. SDTM was developed much earlier in CDISC history and as this data began showing up at the FDA, derived variables and values started creeping in. According to the CDISC SDTM Implementation Guide, a variable is derived when its values "are calculated by an algorithm or reproducible rule, which is dependent upon other data values." Although SDTM was designed to reflect only collected data, several SDTM variables by their nature are derived, while others may be a combination of collected and derived data.

Overall, in order for the models to move towards less subjectivity and inconsistency, the industry must reduce and remove as much derived data from SDTM as possible and maintain this information within ADaM.

LARGE OR SMALL

Another question and one that is probably much more debatable regards how much data to squeeze into a single ADaM data set. The ADaM Implementation leaves that up to the industry and everyone seems to take a different approach. The Basic Data Structure within ADaM supports the ability to add as many parameters as you would want but should you? For example, if you have significant questionnaire data, should you create one ADQS with many parameters and large data sets, or smaller ADaM data sets for each questionnaire? While debatable, our strategy is usually to recommend creating smaller data sets grouping one or a few parameters together that have identical analyses. This limits the size of the data you have to work with and also reduces the amount and complexity of the metadata you have to create for a single data set. In addition, not only should you determine how to divide your ADaM data sets but you should define rules and controlled terminology for the data set names and metadata.

SOURCE OF THE ANALYSIS

Another discussion which often happens within an organization is what to use as the source for your Tables, Listings, and Figures that are created for your study report. In most cases, people say "It depends" or "some SDTM, some ADaM and some raw data" or "Listings come from SDTM and Tables from ADaM". Again, this inconsistent and nebulous approach is about as far from standards as one can get. Once again, while this is debatable, in order to put rigorous processes and objective standards in place, all analyses defined and planned within the study protocol and statistical analysis plan should be created from the ADaM data sets. This will ensure a complete set of analysis data, reduce the subjectivity of deciding what analyses should be created from SDTM versus ADaM, and reduce the amount of data manipulation that occurs within the program code.

INTEGRATION

PhUSE 2011

The final and probably biggest challenge in developing your ADaM standards and data within your organization's overall data flow is determining how the data should be integrated within a submission to generate ISS and ISEs. This decision very much depends on how many differences exist in your SDTM data, ADaM data, and the analyses performed across the studies. The general approach which has been the most successful is to build an integrated SDTM database focused on reconciling terminology and other content issues across your studies. Then, integrated ADaM specifications can be developed by reviewing across the individual study ADaM specifications and adjusted where required. This process also makes defining and following the traceability easier than attempting to combine ADaM data sets and metadata.

DEFINING A STANDARDS GOVERNANCE STRUCTURE

The sections above described the steps and process for developing analysis data elements starting with an analysis result and building a ADaM library to support that development as well as opinions around a successful data flow. However, this process was described in a vacuum without an understanding of how this development and the associated outputs are managed. While this process is good for the initial development of your standards across studies, a overarching governance process must be piloted and implemented to ensure successful implementation, enforcement, and use of the standards across the organization.

WHAT IS GOVERNANCE?

Governance is one of those words which can sometimes carry a negative connotation as a lofty word that really doesn't mean anything. However, if implemented with concrete steps, it can be extremely beneficial to an organization and provide significant efficiencies. Standards governance is the management framework within which standards are developed, tested, and managed. Therefore, the role of standards governance is to provide a decision making framework that is logical, robust, and repeatable to govern an organization's development and use of standards. There are three major components to a standards governance structure within an organization as shown in Figure 2 below.

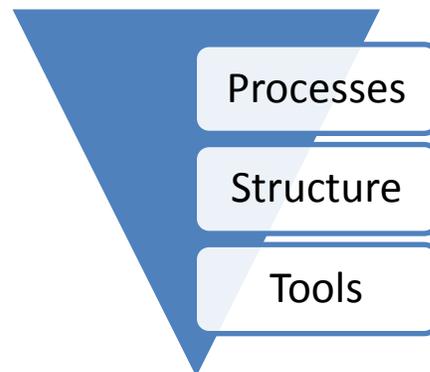


Figure 2. Standards Governance Structure

PROCESSES

The largest, most important, and often overlooked component of a governance structure is the processes and the absolute necessity to enforce those processes. More often than not, companies focus on creating teams of people to sit around and pontificate about data standards and dream up the tools to support them. However, they fail to put the necessary effort into the process of developing, implementing, and maintaining the standards. The following sections outline the different processes that need to be developed as part of a standards governance infrastructure as it relates to ADaM.

ADaM Development

The first step is to implement a process to support the development of the ADaM components. The sections above outlined a process that focuses on an analysis result and develops the analysis elements required to generate that result. In addition to a process for defining a single analysis result you the broader process must be developed for defining, designing, and approving ADaM elements across studies, compounds, and therapeutics areas. This process might include the following steps:

- Identify new Analysis Result
- Define required ADaM elements including ADaM variables, terminology, value level metadata, and derivations
- Submit the ADaM elements for review and approval
- Modify ADaM elements if needed

PhUSE 2011

- Approve and add to the ADaM library

Another key process is to determine the overall standards structure and at what level and scope standards will be created. For example, within one company we tagged ADaM elements at one of four levels:

- Global – Same definition across the organization
- Therapeutic – Same definition across a therapeutic area
- Submission – Same definition across a submission
- Study – Unique to a study

However, this simple definition has many complexities to it, so we defined an additional scope for each of the granular ADaM elements including variable, terminology, parameter value, and derivations. The table below provided an example of different scope definitions.

Variable/ Parameter	Variable Scope	Terminology Scope	Parameter Scope	Derivation Scope
RACE	Global	Global		Global
ITTFL	Global	Global		Study
TRT01P	Global	Submission		Study
PARAM: Tumor Response	Global	Global	Therapeutic	Therapeutic

Table 7. Scope Definitions

In the example above, the scope of the variable RACE is Global as well as its associated terminology and derivation. In the case of TRT01P, the variable name has a global scope, but the terminology will be managed within a submission and the derivation defined at an individual study level. This matrix method of defining the scope of each ADaM element allows flexibility in defining and managing the different pieces.

Implementing ADaM Standards

Another integral set of processes needed are those that must support the implementation of the standards. This includes giving the users of the ADaM standards very comprehensive and easy to understand steps for using and adding to the standards. Users who will need to develop new ADaM elements must understand:

- How they will define an analysis result?
- What different types of elements needed to support the analysis?
- How they access the ADaM library?
- How do they add to the ADaM library?
- Who reviews and approves new ADaM elements?

Users who will implement the standards need to understand where and how they access the ADaM library, how they extract ADaM elements for their study, and how they communicate gaps where ADaM elements have not been defined for their study's needs.

Maintaining ADaM Library

The final processes that must be developed are probably the most challenging. They include the ongoing maintenance of the ADaM library. The challenges that exist in defining these processes are primarily due to the lack of tools available for managing metadata. In general, users need the ability to store and extract the ADaM elements control versioning at the most granular level, links the different ADaM components, and provide ease of modifying, extending, adding and retiring individual ADaM elements. Currently, in today's environment, this is where managing the ADaM elements becomes a very manual and cumbersome process. The lack of adequate tools is the key roadblock in supporting this process.

PEOPLE

The second most important component of a standards governance process is to build an infrastructure of the right people to support both the development and maintenance of standards. While this should not require an army of standards experts, it does require dedicated resources who have the ability to bring different types of users together and are given the authority to enforce the standards and associated processes. The key to not overloading a few key individuals with defining all the standards is to spread out the responsibility across the different standards levels.

Figure 3 shows an example of how a structure could be put in place to support this effort.

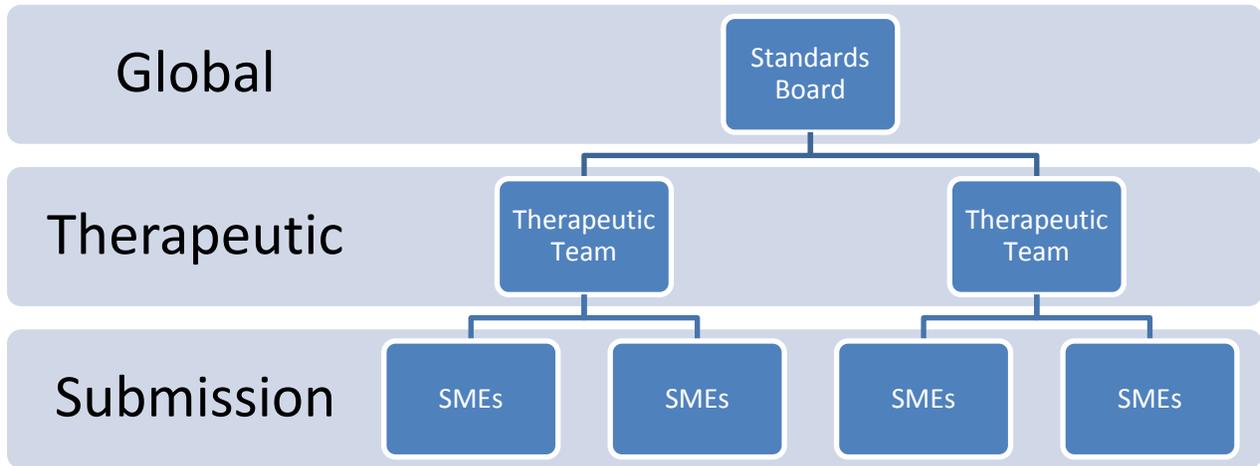


Figure 3. Standards Governance Infrastructure

The table below provides an example of who might sit on these teams and what their responsibilities might include.

What?	Who?	Responsibilities
Global Standards Board	Includes combination of permanent standards experts and representatives from each therapeutic area	<ul style="list-style-type: none"> Enforcement of standards across the entire organization Development and maintenance of global standards Ensure standards alignment and controlled terminology across therapeutic areas Ongoing alignment with industry standards
Therapeutic Teams	Team consisting of one or more people from the indications/phases within the therapeutic area; Team lead would sit on the Global Standards Board	<ul style="list-style-type: none"> Enforcement of standards across the therapeutic area Development and maintenance of therapeutic area standards
Submission SMEs	One or more subject matter experts would be defined for each submission and meet regularly with the therapeutic standards lead	<ul style="list-style-type: none"> Provide subject matter expertise to develop specific ADaM elements Review and approve all submission specific ADaM elements

Table 8. Responsibilities of Standards Team

TOOLS

The final component to make standards governance successful is the implementation of tools to support the management of ADaM metadata. While I know everyone reading this paper has waited anxiously to hear about the great technology solution that will make all of this automated, but unfortunately I will fail to provide the proverbial Holy Grail.

The industry has struggled for many years to implement the right tools to support the standards development process and the fact is that no real tools for managing metadata exist. In lieu of robust solution, the industry continues to use, or misuse tools such as Excel for developing and managing standards. The cartoon below, while humorous provides a realistic perspective on how people misuse tools to attempt to manage standards.

PhUSE 2011

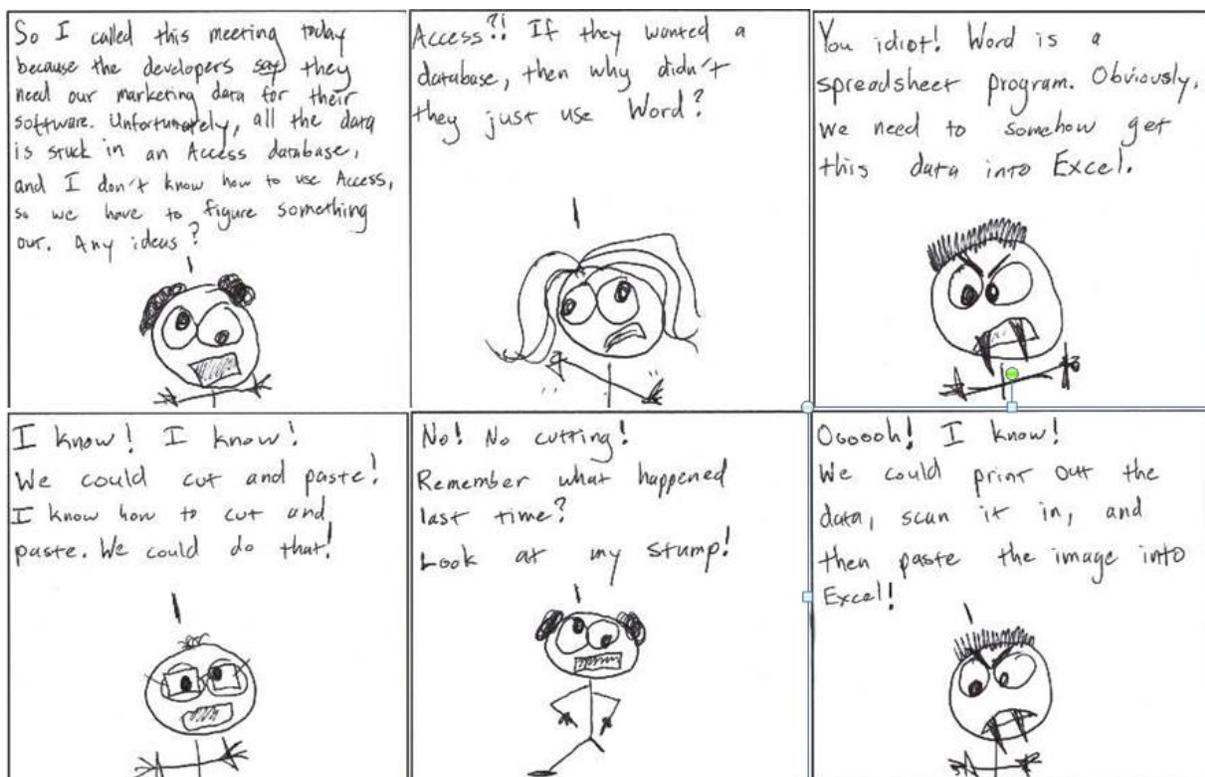


Figure 4. Cartoon regarding the incorrect use of tools

The biggest issue that we face in the industry is that no tools exist that focus solely on the need to manage metadata. All the tools we do use for processing clinical data from EDC to ETL to data repositories focus on a specific core functionality with metadata as an afterthought. So for now, we have to live with Excel until vendors create solutions that can handle the development, implementation, and maintenance of metadata.

BEST PRACTICES/LESSONS LEARNED

While a number of best practices and lessons learned have been intertwined within this paper, this section highlights a few specific topics to consider when implementing data standards.

STOP CODING

The first thing our industry always has a tendency to do is immediately begin coding or in this case, mapping clinical data. We map and code, map and code, and without realizing it, we continue to reinvent the wheel generating a plethora of uncontrolled and unmanageable mapping files. In order to effectively create a robust standards governance infrastructure to support the development and implementation of standards, your organization must stop coding, step back, and develop well defined and tested processes and tools to support the process. The biggest challenge of stepping back and focusing on a robust process is the need for people, money, and time, all of which are hard to come by within organizations. However, if you don't take the time to develop this foundation, you will never realize any efficiencies in the development and use of standards.

AVOIDING "IT DEPENDS"

One of the biggest advantages of implementing ADaM is its flexibility to support a wide range of analyses. Unfortunately, that advantage is also its greatest weakness. In numerous conversations with customers and within the CDISC community, when asked to make a decision regarding an ADaM naming convention, algorithm, or subjective guideline, the answer is always "it depends". Unfortunately, the answer "it depends" means you will never fully implement an analysis data standard and will fail to make your organization more efficient. Wiggle room must be removed, subjective decisions must be made objective and, whenever possible, remove the grey areas. In most cases it's not the analysis that is unique, but the people implementing the analysis.

CONTROL, CONTROL, CONTROL

The most common failure of standards management is the lack of control. Companies attempt to define standards but don't really develop a governance process that tightly controls the development and management of the standards. Part of the issue is the lack of tools to support the process, but another large part of the issue is the flexibility which companies allow as described in the previous section. It is critical that the company define and maintain a metadata library no matter how manual the processes are in the beginning. This library should be managed across all ADaM elements and all levels of your standards (e.g. Global, Therapeutic, Submission, and

PhUSE 2011

Study). This might seem cumbersome and without fancy tools, it can be very time consuming. However, without this level of management and control, the pain downstream will greatly outweigh early challenges.

SMALL BITS AND ITERATE

Experience has shown that companies will try to tackle the entire standards process in one huge gulp, and in most cases, they initiate many different tasks and end up not doing any of them every well. You cannot solve the standards development, implementation process, and tools all at once, especially within a larger organization. The goal should be to identify small bits first and work on developing and testing the process around those bits. For example, within the ADAM development process, the first step might be to look across three critical therapeutic areas and focus on developing core ADSL elements. At the end of this exercise you might only have a dozen or so ADSL elements, but those elements are well defined and tested. The next step would be to iterate through additional bits of ADaM elements expanding across different analyses and therapeutic areas. Developing standards and governance using this iterative methodology will allow your company to refine the process as well as continue to add to your overall metadata library.

METADATA GONE WILD

In this last section, let's actually play devil's advocate, and tell you to stop collecting mountains of metadata. While the collection and use of metadata as described in this paper sounds like a great idea, in practice, this can be a very time consuming and nearly impossible workflow. This is especially the case when you don't have the tools to support metadata collection. The last thing you want to do is collect metadata just to say you did it and not be able to use that metadata to improve your development process. It's the time tested analogy of the chicken and the egg. Given this challenge, we recommend tackling the issue of collecting and using this metadata in small encapsulated pieces. The goal is to start small, refine the process as you go along, and look towards developing tools to support metadata management, or collaborate with vendors to develop solutions that will support this need to manage this metadata.

CONCLUSION

Implementing standards within an organization is much more than creating a spreadsheet and entering some metadata, and this is even more evident when defining analysis standards. These standards include so many moving parts including the data, variables, derivations, and complex statistical methods, all of which needs to be captured within a well defined framework. ADaM is new and the industry is still in the beginning stages of adopting it for analysis data. Your organization's goal should be to start small, make sure you understand the development process, refine as you iterate through the process, and build up your ADaM library with all the components described within this paper.

The standards implementation process must include three key components to be successful: an agreed upon data flow, a set of robust standard elements that don't include "it depends", and a well defined and enforced governance process. Making sure you implement all three components successfully can be challenging, but if you can, your organization will realize significant efficiencies throughout the process.

CONTACT INFORMATION

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

Chris Decker
d-Wise Technologies
4020 Westchase Blvd, Suite 527
Raleigh, NC, 27519 USA
Work Phone: +1 919 600 6234
Email: cdecker@d-wise.com
Web: www.d-wise.com

Brand and product names are trademarks of their respective companies.