

## PhUSE 2014

Paper CD07

# Is Your Clinical Data Standards Library Ready for Change? – Adapting to Therapeutic Area Specific Standards

Author: Lauren Shinaberry, Business & Decision Life Sciences, Brussels, Belgium

## ABSTRACT

The release of Therapeutic Area User Guides (TAUG) based on priorities provided by CFAST is an ongoing project; among the first user guides released are Asthma, Diabetes and Alzheimer's. The TAUG are critical as they give guidance on mapping therapeutic specific key data to SDTM. It is important to understand the compatibility of your current standard library with the new TAUGs. Implementing standard libraries in an integrated metadata management tool enables you to have an overview of your libraries and manage maintenance.

This paper provides an overview of the current, ongoing and planned TAUGs from CFAST. A description of the process of the TA standards development is presented as this supports the discussion of the approaches to organizing your global and TA standards libraries. A brief discussion of the historical progression of methods for storing metadata is given as well.

## INTRODUCTION

With multiple and frequent releases of TAUG planned by CFAST it is necessary to develop a flexible and robust approach to creating and updating TAUG standards. Traditional standards libraries have often included a single standard containing all expected versions or large arbitrary separations of standards into early phase standards and late phase standards. The benefit of traditional standards library models was the reduction in administration around naming conventions and updates. As libraries evolved and integrated into large sized pharma companies Therapeutic Area / Disease Specific libraries were created and maintained. The updates of these can take 6 months to 1 year review and update cycles depending upon governance processes and platforms. This allowed libraries to adapt to internal changes in systems and sponsor specific data collection and analysis end points.

With the release of the draft guidance (February 2014) FDA Standardized Study Data for Electronic Submissions a rapid and external driver of standards change can be expected. Adopting to the forecasted changes early and using an automated approach to metadata version management allows the sponsor to take ownership of the standards benefits from the updates while not necessarily increasing the administrative burden of traditional and integrated Therapeutic Area libraries.

The CFAST list of planned TA standards and their release dates can be found on the CDISC website: <http://www.cdisc.org/therapeutic> and an example from September 2014 is shown below.



## Program Overview – September 2014 Approved Therapeutic Area Standards Projects

Therapeutic Area	Coordinating Organization(s)	Proposal Approval Date	Stage 0	Stage 1	Stage 2	Stage 3a	*Stage 3b	*Stage 3c
	Project Manager		Scoping & Planning	Concept Modeling	Standards Development	Internal Review	Public Review	**Projected Publication
CV Endpoints v1	CDISC/DCRI Amy Palmer	Jun 13	July	Sep	Nov	Feb	May	Q314
Diabetes v1	TCB Rachael Zirkle	Apr 13	May	Aug	Dec	Apr	May	Q314
QT Studies v1	TCB John Owen	Aug 13	Oct	Feb	Mar	Jul	Sept	Q414
Traumatic Brain Injury v1	CDISC Rhonda Facile	Oct 13	Sept	Oct	Nov			Q215
Chronic Hepatitis C Virus v1	TCB John Owen	Nov 13	Feb	Apr	Jul	Oct	Nov	Q414
Schizophrenia v1	CDISC/DCRI Amy Palmer	Nov 13	May	Jul	Aug	Oct	Nov	Q115
Breast Cancer v1	TCB Pam Harvey	Nov 13	Aug	Sept	Nov			2015
Influenza v1	C-PATH Laura Butte	Feb 14	May	Jul	Aug	Sept	Oct	Q414
Dyslipidemia v1	TCB John Glover	Dec 13	May	Sept	Oct			Q115
COPD v1	TCB John Glover	Nov 13	Aug	Oct	Dec			Q315

Key |   Stage completed |   Stage ongoing | All months reflect when stage is, or is projected to be, completed.

\*The Stage3b concludes at the end of the 30-day review period and Stage 3c concludes when all tasks have been completed and the standard is publically available.  
\*\* Specific projected publication dates to be added to the notes section at the conclusion of Stage 3b.

September 19, 2014



### Definition of content

Development of TA standard content requires an understanding of the clinical aspects of the data that is unique to the area. The approach that has been developed by CDISC to facilitate the creation of TA standards as part of the CFAST project is a good model to consider in the development of internal TA standards.

A critical feature of the CDISC-CFAST process is starting with the definition of the scope and clinical concepts. Through the use of mind-maps, the development of the standards begins by clearly understanding the information that is collected as well as the clinical processes involved. Rather than starting with the SDTM data, the concept mapping approach ensures that a clear definition of the data and its use in clinical research underlies SDTM mapping decisions. Below is an example of a mind-map from the Alzheimer's TAUG.

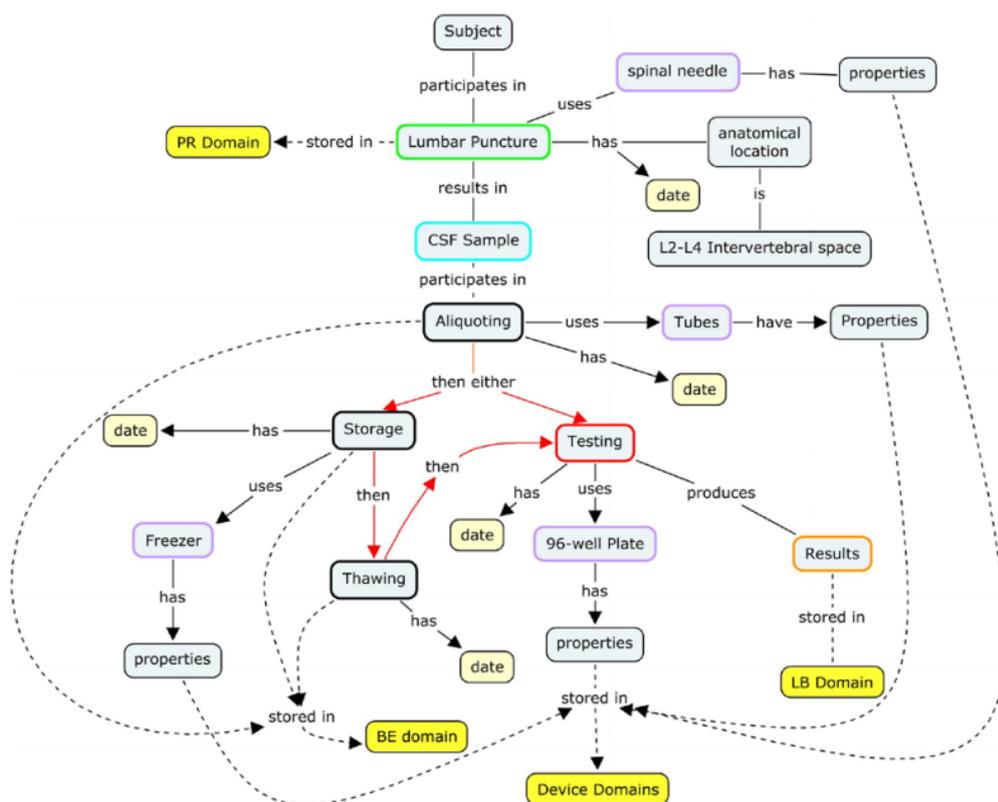


Figure 3.1: CSF Sampling and Processing

While the mind-map or concept map may look complicated at first, it has the advantage of ensuring that a complete picture of the types of data that will need to be represented is in place prior to the process of mapping it to standards such as CDASH, SDTM and ADaM. This is an improvement over previous methods that often relied on representative examples from several clinical studies that may not always have given the complete context.

The CDISC-CFAST TA development process helps identify those aspects of the data that may require updates to the standards in cases where the complexity of the TA does not fit the existing models. However, changes to the foundational CDISC standards would only be made if the existing models were insufficient. Any updates made to accommodate TA data will be added into the foundational models as well. This is important from the perspective of managing your metadata in that the TA standards and the foundational (e.g. global) standards are not mutually exclusive.

## Managing content

Once the content of the TA standards is defined, it needs to be managed and made available to the end user community. There are two approaches that will be explored in this paper:

- Fully incorporated: Separate standalone libraries that incorporate all globally applicable standards with the TA-specific standards
- Modular: A single global library that is complemented with discrete TA-specific libraries

The benefit of fully incorporated libraries that contain both the global standards plus the TA standards in a single library is that it simplifies the end-user experience. There is a single place to reference for the standards that are needed. However, this can introduce difficulties for the management of the libraries. By duplicating the global content, the risk of inconsistencies is high and effort to make updates to global metadata is increased.

The modular approach of having a single global library that is referenced by all TA user communities in combination with referencing another library for the TA-specific metadata makes it more efficient for librarians to manage updates in one central repository. However, the end users may have more difficulty knowing how to use multiple libraries.

The ideal solution would allow the independent libraries proposed in the second approach for ease of governance and management, but present the metadata to the end users in a consolidated "single library" view.

## Methods for storing the metadata

The traditional method for storing the definition of data standards has been Excel spreadsheets or other document-based processes. These may be stored in a file server environment or through a versioned document management system. These systems are widely available and well known to most users reducing the start-up times and trainings associated. Some drawbacks of this method are that it is a two-dimensional representation and requires manual effort to maintain. Additionally, it can be difficult for end-users to understand how to implement the standards and the version control is usually a manual process.

Domain Prefix	Variable Name	Variable Label	Type	Controlled Terms, Codelist, or Format	Role
AE	AETERM	Reported Term for the Adverse Event	Char		Topic
AE	AESEV	Severity/Intensity	Char	AESEV	Record Qualifier
AE	AEREL	Causality	Char	*	Record Qualifier
AE	AEDOUT	Outcome of Adverse Event	Char	OUT	Record Qualifier

Figure 1: Example spreadsheet representation

In recent years, the implementation of metadata repositories based on relational databases allows more robust governance abilities. From the viewpoint of therapeutic area libraries, relational database platforms offer the ability to combine the benefits of both the modular approach that is more manageable from the librarians' point of view with a dynamic, linked view of the libraries in the manner of a fully incorporated library.

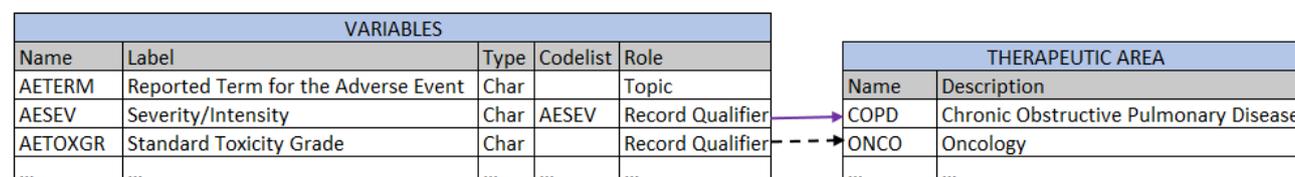


Figure 2: Example Relational Database (RDB) representation

In Figure 2, the arrows between the VARIABLES table and the THERAPEUTIC AREA table are indicating how you could indicate which variables are specific to a particular TA. The solid arrow is indicating AESEV is used in the COPD library, while the dotted arrow indicates that AETOXGR is used in the ONCO library. The AETERM variable would be considered part of the global library as (in this example) there would not be any TA-specific implementation of the variable – both example TAs would use it in the same way. This avoids duplicating information in multiple locations within the database.

The additional advantage to the relational database/metadata repository platform is that it facilitates easily identifying where TA-standards differ from the global standards and can present these in a hierarchical manner to the end users. In other words, the TA-implementation of the standards can be presented as the primary choice, with the more general global implementation offered as a secondary option. The drawbacks associated with these systems can be the implementation time and trainings associated.

Taking the modular concept one step further, we are beginning to see the use of classic conceptual modelling approaches that are likely familiar to people with IT or computer science backgrounds, but are only recently

being spoken about in the clinical data standards end-user community. Not too many years ago, the term “metadata” was a new concept to most end-users. These days, terms such as “semantic interoperability” and “ontologies” are replacing “metadata” as the cutting edge terms in clinical data standards implementation. These are not new concepts, but they may be relatively new to clinical data managers and librarians.

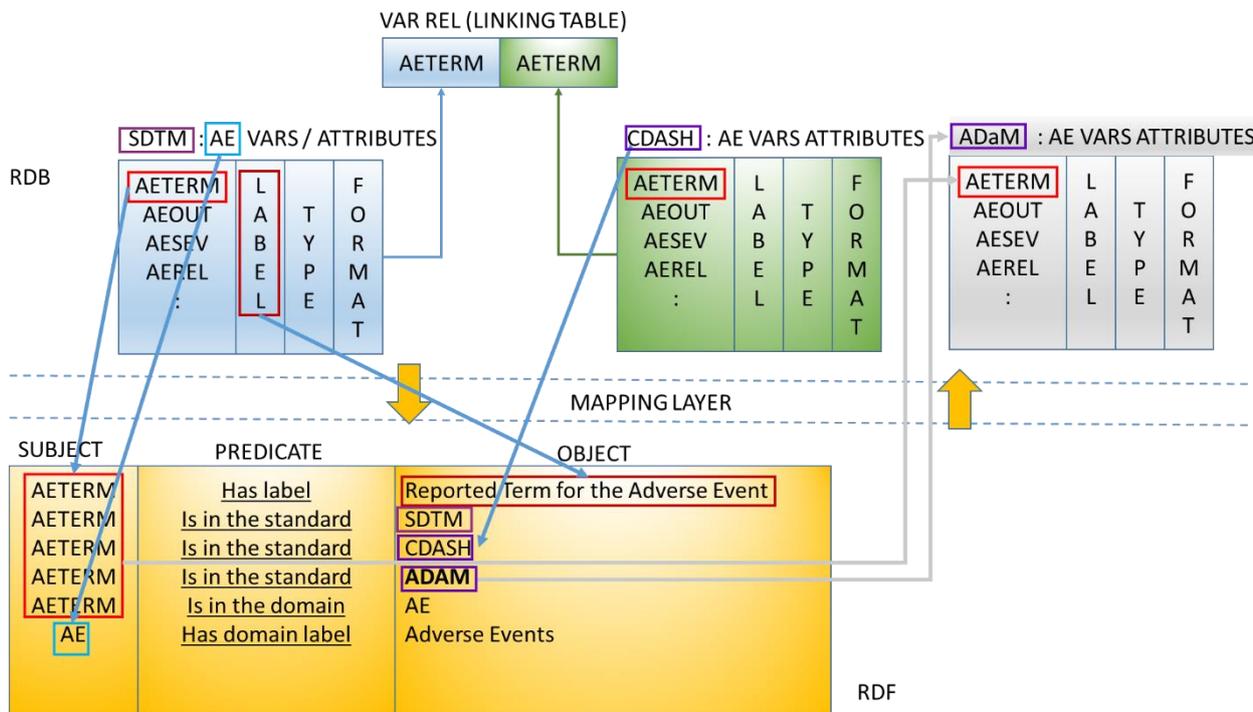


Figure 3: Conceptual example of Resource Description Framework (RDF) representation

For more information on the use of RDF with CDISC standards, there are numerous posts on the PhUSE Wiki and PhUSE website. The relevance to the topic within this paper is that it is the next logical progression for managing the increasingly complex and interrelated nature of therapeutic area standards.

## Conclusion

Therapeutic area standards are not a replacement for globally-applicable standards. Many of the safety domains may be consistent across many different TAs and the TA-standards should be seen as complementary implementations of the global metadata. The optimum method for managing the metadata of both global and TA standards would mirror this relationship.

As an industry we have been evolving from the two-dimensional approach of Excel files to the more robust use of relational database that can support providing efficient metadata management without duplication of content. The next logical step is to move into methods of representing ontologies that can support end-to-end integration. It is important to balance the needs of the librarians with the needs of the end users. The most elegantly designed metadata repository is meaningless if the content is not easily understood and accessible to the user community. So, regardless of the method behind the scenes, the content must be easy to use.

The CDISC-CFAST development of TA-standards mirrors the modular approach to library management. CDISC TA-standards are designed to supplement the foundational standards and do not replace them. TA standard implementation guides only include those implementations that are specific to the area, rather than duplicating the standard domains that are not TA-specific. It may be possible that new CDISC metadata may be introduced to support TA standards if the existing foundational standards are insufficient, however these new standards will be incorporated to the foundational implementation guides.

## CONTACT INFORMATION

Contact the author at:

Lauren Shinaberry

Business & Decision Life Sciences

St Lambertusstraat 141 Rue Saint-Lambert

1200 Brussel – Bruxelles

Work Phone: +32 2 774 11 00

Fax: +32 2 774 11 99

Email: [Lauren.Shinaberry@businessdecision.com](mailto:Lauren.Shinaberry@businessdecision.com)

Web: [www.businessdecision-lifesciences.com](http://www.businessdecision-lifesciences.com)