

Developing MDR Metadata Requirements for Operational Implementation

Andy Richardson, d-Wise, Manchester, UK.
Scott Bahlavonni, d-Wise, Morrisville, NC, USA.

ABSTRACT

Metadata Repositories are increasingly being turned to as the solution to the management and operational implementation of the CDISC and other standards. In order to be successful, this requires a MDR to hold, integrate and function both as a library of defined standards and as an operational tool able to support additional requirements such as version control, specification generation, data traceability and compliance. This requires users to provide, in addition to functional requirements, details of all of the specific metadata defined content to support their operations. Using standard operational models and a systematic metadata categorisation methodology, this paper will illustrate how a complete metadata defined MDR specification can be developed and tested to support the evaluation, implementation and integration of a MDR solution to support day-2-day standards-driven clinical data operations.

INTRODUCTION

The need to improve efficiencies, maintain high quality and meet regulatory submission standards in clinical data is driving a fundamental re-think in how clinical data systems are designed and operated. Metadata Repositories are now considered a fundamental part of these improvements, but require additional user specification for successful implementation, specifically to define exactly what metadata is needed to support the operation.

In this poster we extend previous work in this area to show how by using a combination of Standards Operational Models and Metadata Classification, vendor independent MDR definitions and content can be systematically developed to support successful MDR operational implementation.

4-STEP METADATA DEFINITION METHOD

Using the four step methods described below a comprehensive set of metadata definitions can be developed to support all operational requirements over the clinical study lifecycle.

Step-1: Identify the Entities/Artefacts & Relationships required by the operation (using, for example, Standards Operational Modelling).

Step-2: Identify the general and specific sets of Metadata Categories required to define all the attributes required to manage, govern, specify and control each Entity/Artefact within the operation.

Step-3: For each Entity/Artefact * Metadata Category develop the specific metadata required by each combination.

Step-4: Using test data and examples, confirm the metadata definitions

The resulting *Vendor Independent Operational Metadata Specification Products* are then available to support Use Case requirements, the evaluation of specific MDR solutions, as containers to build specific content in preparation for loading into an MDR, and to develop User Acceptance Tests.

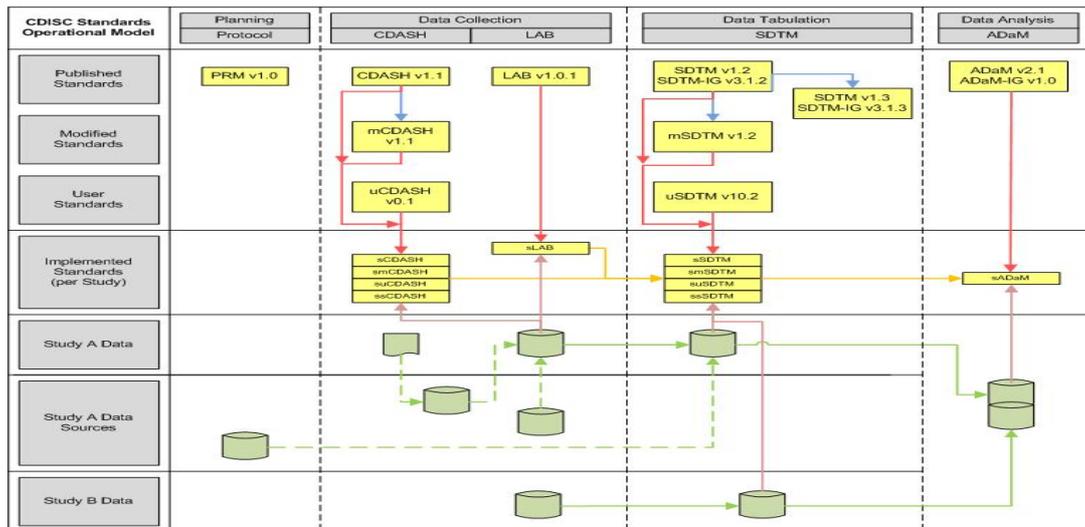


Figure 1: An example Standards Operational Model showing the details of the entities required over the clinical study lifecycle, and their relationships.

1: ARTEFACT & RELATIONSHIP IDENTIFICATION

In previous work [1] we have shown through the use of Standards' Operational Models how the details of all the objects (artefacts) required to define and operationally implement CDISC or other standards throughout the clinical data operations can be systematically developed.

Figure 1 (above) is a simple example of a standards operational model and shows which standards are used by the organisation over the study lifecycle (top row), and the various states they exist in in the organisation (left column). There are 40 potential standards-by-state combinations in the diagram, of which 14 (in yellow) are required to define or specify a study's data, and 9 (in green) are the standards implemented as study datasets.

Having identified which artefacts are needed, the links (relationships) between them can be added. These also will require to be recognised within an MDR; the most important of these being the mapping relationships required to convert clinical data between the various standards.

All of these objects can be described to some degree in metadata terms; ranging from the full descriptions of the CDISC standards (first row), through to any specific information about, for example, the specific location of a SAS dataset that is to be managed within an MDR.

Entity	Artefact
:	:
SDTM	SDTM (Model)
SDTM	SDTM IG
SDTM	TAUG
SDTM	SDTM Domain
SDTM	SDTM Domain Variables
:	:
Controlled Terminology	Codelists
Controlled Terminology	Codes
:	:
ADaM	ADaM (Model)
ADaM	ADaM IG
ADaM	ARM
ADaM	Computational Method
ADaM	Dataset
ADaM	Dataset Variables
:	:

Figure 2: Section of a Metadata Artefact Definition document showing the sets of CDISC SDTM, CDISC ADaM and Controlled Terminology artefacts that are required to be defined, and the content managed, using a commercial MDR system.

Step 1 is now to decide precisely which artefacts will be described and managed within the MDR. This is most commonly those in yellow in Figure 1, however the specific details will depend upon considerations, for example, organisational roles and responsibilities, or the management of specific artefacts by other applications (e.g. the data collection objects by an EDC system). The result of this exercise is the list all the artefacts and relationships required to be metadata defined (Figure 2).

2: METADATA CATEGORISATION

Whilst the principal requirement for the clinical data operations MDR is to manage the metadata definitions describing the CONTENT of the artefacts identified in Step 1, this is only part of the metadata information required to support all the operational requirements. For example, the standards management groups will require an MDR to support standards ADMINISTRATION (or governance).

Unlike CONTENT definitions, which are artefact specific, these operational metadata are required by *all artefacts* and will form the basis for integrating the MDR CONTENT definitions into the clinical data operations. Once identified and described, combining these 'categorisation blocks' enables a full description of all the information required to describe and manage each artefact within an MDR.

[StandardsModel]				Content					Administrative			Operational		
sm1	sm2	sm3	sm4	c1	c2	c3	c4	c5	a1	a2	a3	o1	o2	o3
...
...

Figure 3: Overall shape of a metadata category defined artefact. The colour coding shows how the categorised metadata elements build into a complete description of the artefact.

Metadata Categories and Subcategories

CONTENT Metadata

The "Content" metadata category hold the metadata necessary to describe a specific artefact, e.g. a SDTM domain, and will be unique for each artefact. The CDISC eSHARE published standards are examples of externally available "Content" metadata.

ADMINISTRATIVE (or GOVERNANCE) Metadata

"Administrative" metadata consist of the metadata required within the organisation to manage the "Content". Common metadata in this category are "Version", "ValidFrom", "ValidTo", (subcategory: "Content Management"), but may also include metadata such as "Content Owner" or "Content User" (subcategory: "Content Authorisation"). This category is common to all artefacts.

OPERATIONAL Metadata

The 3rd metadata category - "Operational" - adds those elements required to recognise where and for what purpose "Content" is being used. Typical operational metadata are "Study", "Dataset Location", "Compliance Status" and so on. As with "Administrative" metadata, this group is common to all artefacts. Figure 3 presents an overview of the final artefact metadata structure.

The clear recognition and differentiation of the different required operational metadata types enables efficient independent development and review, and can be extended to meet specific organisational requirements.

3: METADATA ELEMENT DEVELOPMENT

Figure 4 shows an example of the final metadata requirements needed for operational implementation, and illustrates the value of the categorisation method.

SDTM-Model		Example Expected MDR Records	
Metadata Category	Metadata Element		
Content	Model_Version	1.4	1.4
Content	Reference_URL	<>	<>
Content	Observation_Class	All Classes	All Classes
Content	Structure		
Content	Variable_Order	2	2
Content	Variable_Name	DOMAIN	DOMAIN
Content	Variable_Label	Domain Abbreviation	Domain Abbreviation
Content	Variable_Type	Char	Char
Content	Variable_Role	Identifier	Identifier
Content	Variable_Description	Two-character...	Two-character...
Administrative	Status	Active	Active
Administrative	ValidFromDate	Aug-20xx	Aug-20xx
Administrative	ValidToDate		
Operational	StudyID		WDYT-091-001

Figure 4: Example final metadata elements describing for the SDTM-Model. The left columns show the detail of the use of metadata categorisation to complete the final definition. The right columns show 2 examples of the expected MDR data illustrating how the method can review and test different operational use cases (here, adopting the model for use by a study)

4: METADATA DEFINITION TESTING

Having developed the specific details for any particular artefact, the definition can be tested by completing example records reflecting the content, administrative or operational metadata required to support any specific use case.

The resulting vendor independent specifications can then be used to support MDR evaluations, use case development and acceptance tests, or to create content for loading into commercial MDRs.

SUMMARY

The 4-Step Method described here for developing operational MDR metadata definitions and content provides a robust, vendor independent, and easily understood approach to creating these important specifications.

The method is particularly powerful in ensuring that the use of standards throughout the clinical study lifecycle is fully understood, and shows unambiguously where and what metadata “consumer” systems (e.g. EDCs, SCEs etc.) need from standards management MDRs.

REFERENCE

Richardson: A Model for Reviewing the Operational Implementation of the CDISC Standards, Presentation CD06, PhUSE, 2015

CONTACT INFORMATION

Author Name: Andy Richardson
Company: d-Wise
Address: Peter House, Oxford St, Manchester M1 5AN, UK.
Work Phone: +44 161 209 3670 - +44 7985 245 416
Email: andy.richardson@d-wise.com
Web: www.d-wise.com

Author Name: Scott Bahlavonni
Company: d-Wise
Address: 1500 Perimeter Park Dr. Suite 150 Morrisville, NC 27560, USA.
Email: scott.bahlavooni@d-wise.com
Web: www.d-wise.com