Paper SI05

Maturing standardisation; the switch from spreadsheet-based standards to an MDR

Louella Schoemacher, OCS Life Sciences, 's-Hertogenbosch, Netherlands
Jasmine Kestemont, Innovion, Putte, Belgium

ABSTRACT

In order to work consistently, structured, and time efficient throughout multiple studies, standardisation is key. Standardisation can be relatively simple with having the company standards in a spreadsheet and using this spreadsheet to share study-specific metadata requirements.

However, with a growing number of compounds, indications, studies, and number of CRO's involved in creation of the SDTM package (aCRF, SDTM datasets, Define-XML and SDRG), it gets harder and more important to have and maintain the standards. Furthermore, individual studies can have specific requirements which may result in deviations even within a compound or indication. When the number of studies grows, a simple spreadsheet may no longer be sufficient to capture all subsets of metadata requirements. When the standards cannot keep up with the growing number of studies, this can result in an overgrowth of different metadata definitions, leading to undesirable inconsistencies.

To prevent overgrowth and to keep as much consistency as possible, a solution is the implementation of a metadata repository (MDR). But how to make this switch from having standards in spreadsheets to an MDR?

First of all, what are the requirements for such a system and how can it be implemented within the organisation? What do we want to get out of an MDR? Is it purely about the storage of metadata or do we want additional features, such as the ability to build an actual industry compliant Define-XML? How does such a system work and does it have to adhere to GxP standards? And following on the GxP discussion: how will the system be validated?

Secondly, how will we govern these processes? Who is going to build the standards? Are we going to have general standards, or will we create indication specific standards? How often will these standards be updated? Are we going to upload the current standards into the system, or will we start from scratch? Will we use the existing CRF questions or are we going to transform them to CDASH questions?

Finally, how will we collaborate regarding the standards? What will the internal review and approval process look like, and which functional areas will take part in this? How will we create study-specific metadata and who will be involved in the set-up? How are we going to collaborate with external parties after implementing the metadata repository and how will it affect their processes?

These considerations are important when setting up a solid standardisation process. This paper will guide you through our most important findings and conclusions so far in our process of implementing an MDR.

INTRODUCTION

In order to work consistently, structured, and time efficient throughout multiple studies, standardisation is key. Standardisation can be relatively simple through company standards in a spreadsheet and by using this spreadsheet to share study-specific metadata requirements.

The use of spreadsheets is straightforward and easily maintainable when only used by a small number of people and in a small number of studies with similar endpoints. However, within a growing organisation, in terms of both employees and portfolio (figure 1), having spreadsheet-based standards becomes a challenge. Spreadsheets are often editable in order to make it study specific, but this results in the files being prone to mistakes. Moreover, it is a challenge to keep track of different versions and to track the origin of a specific item. Another challenge within a growing organisation is to keep oversight of what is collected throughout the portfolio and to note, often small, changes applied to CRFs or Define-XMLs across different studies or indications.

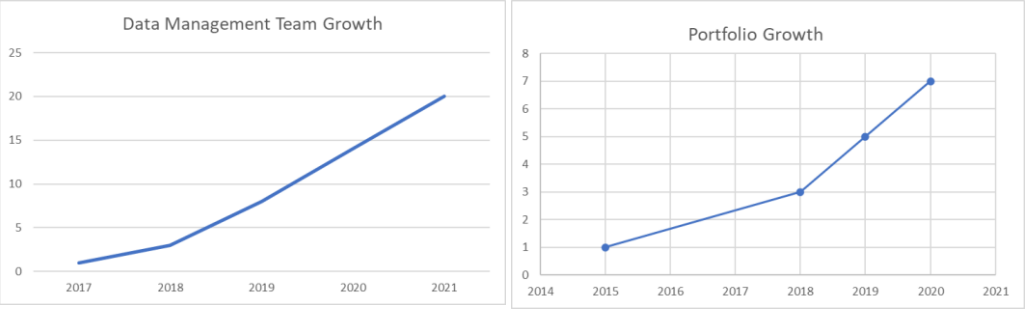

Figure 1. Growth of both data management team and portfolio over the past years.

To prevent overgrowth and to keep as much consistency as possible, we implemented a metadata repository (MDR) over the past year. This paper describes our experiences in the different phases of the implementation of an MDR and where we currently are in the process.

MDR SET-UP PROCESS

SYSTEM REQUIREMENTS

To find an MDR that would suit our needs, we had to think about the minimum requirements for such a system. Firstly, the software should be off-the-shelf and intuitive to use in order to start the actual build of standards on a short notice. Furthermore, content should be easy to re-use, as one of the purposes of implementing the MDR was to work time-efficiently throughout multiple studies. In order to do this, we should be able to re-use forms, codelists and metadata easily throughout the system.

The software should also be able to provide front-end and back-end visualisations. Front-end visualisations to show what the standards would look like in an EDC system, and back-end visualisations to show what the metadata of the standards would look like. Finally, we should be able to version standards. Without having a pre-defined versioning process in mind, the software should at least enable us to use different versions of standards when the SDTM Implementation Guide or the NCI Controlled Terminology is updated.

END USER OBJECTIVES

Next to the minimum requirements of the software itself, we also formulated objectives regarding how the implementation of an MDR should improve processes within the organisation. One of these end-user objectives is to have data standards governance in place. In a fast-growing organisation, it can be easy to focus on ones own study or indication and lose track of what others are doing. However, implementing an MDR forces you to take a step back and to create an overview of what data is currently collected in studies over different compounds and indications, and why these data are collected. This should help us in harmonising CRFs, and eventually collecting more uniform data.

One of the main end-user objectives of MDR implementation is to be able to design standards and use these standards to build study CRFs and corresponding metadata in parallel with protocol development. This way each decision in the protocol is reflected almost immediately on the CRF and issues or unmet expectations can be identified before the protocol finalisation. This would enable us to have an early definition of trial-level expectations, preferably resulting in a smoother study set-up and study conduct.Besides the objectives focused on internal processes, another objective of the MDR is also to have standardised communication with external parties. By developing and sharing standards, we aim to set clear expectations at both sponsor and vendor level, ultimately resulting in improved communication. We hope that the use of standards is not only beneficial to our internal processes, but also to the vendor processes. For example, to simplify user acceptance testing (UAT) of the EDC. Once a form is tested within the EDC, it should not be tested extensively again for another study reducing the time spent on UAT during study set-up.

SELECTION CRITERIA

With these end-user objectives and system requirements defined, the search for a suitable MDR started in March 2021. Different suppliers were selected to show their software. In April 2021, they were requested to present a business case to show if the product could meet the requirements (figure 2).

The criteria used to select a vendor were: the ability to create CRFs, metadata specifications and Define-XML, customisation of CRFs, to govern the standards within the system, cost, and finally a short time to implementation.

In May 2021, a supplier was selected and we got approval from the organisation to start the implementation. The implementation team was led by the project manager and supported by the business owner, application owner, computer system validation representative, quality assurance, and business information systems/IT. The goal here was to have a small, focused team yet including all key decision makers in order to shorten the time to implementation.

After approval and risk assessments of the respective departments, we had to start developing detailed user requirements, elaborating on the requirements and selection criteria. With these user requirements, the development of user acceptance test scripts could start. This was a first step to getting to know the software in detail. By the end of October, the test scripts were fully executed, and the system passed the UAT. Hereafter, the development of our standards could start, fulfilling the last selection criterion: a relatively short time to implementation.

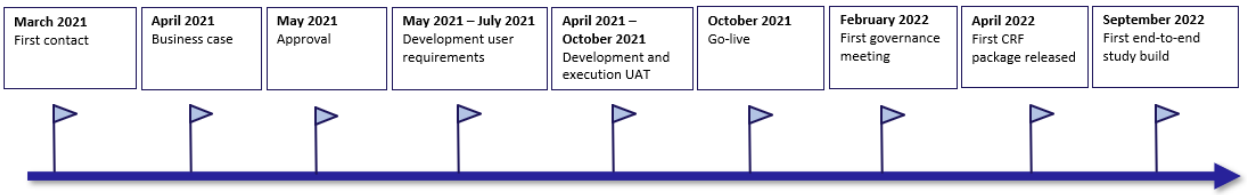

Figure 2. Implementation timelines

MDR SET-UP PROCESS

STANDARDS DEVELOPMENT

As we operate in a rare disease area, we study various indications and different endpoints, for which it can be challenging to develop standards. Because of this diversity, we decided to start building the standards from scratch rather than modifying existing ones. The aim was to create different types of deliverables, however the approach that was taken was to first focus on building CRFs. In this way the focus was first on what data actually needed to be collected, before focusing on instructions (CRF completion guidelines), edit checks and metadata of the collected fields (Define-XML).

To have a representative overview, CRFs of different indications from the past three years are compared to find common items and identify gaps. For each item we reconsider its purpose (operational field versus data collection field), challenge the need for the item to be on the CRF (is it necessary or just nice to have?), try to identify possible challenges the site might have while filling out the forms to overcome data issues at data entry level, and finally check the annotations to have a correct mapping to SDTM.

When making the comparison, we also check the items which are different across the CRFs. Are they different because they serve a different purpose? Are they different because the wording of the question is different? Are they different because the CRF is designed by a different vendor? Or are they different because the indication requires us to collect different information? Based on the all considerations above, we decide if the CRF will be designed for general use, or if it will be designed specifically per indication.

DEVELOPMENT CHALLENGES

As expected, when working with new software, we face some challenges. To start, we sometimes do not know what the software can or cannot do. This means that the system's abilities do not always match up with what we, as users, would expect to be most convenient.

Secondly, as opposed to spreadsheet-based standards, there is not only a learning curve with regards to the standards but also to the software itself. Although the system is quite intuitive and the concept is easy to grasp, there are many features and options, which results in a learning process for the software as well.

A third challenge is how to best apply versioning in the system. When should a standard be released? And should it be released as a whole, or should finalised CRF forms be released separately? Next to setting up a versioning process, this also results in a daily confrontation with the 80/20 rule. Do we wait until the form is perfect, or do we release it when we think it is nearly perfect?

Finally, a challenge arises when starting to add annotations to the forms regarding codelists. The best example to illustrate this is the NCI Controlled Terminology codelist containing the codes “Yes” and “No”. These codes are used across various forms, but have different meanings. On the Eligibility form, the answers have an operational purpose to trigger subsequent questions. On the Adverse Events form, the answer is captured as “Y” or “N” in the variable AESER (Serious Adverse Event), while on the Vital Signs form, there is even a difference between how answers “Yes” and “No” are being mapped (figure 3). Preferably we would like to attach our annotations to the answer options. However, since this codelist serves many purposes, we now attach the annotations to the question rather than the codelist item and manually place the annotations next to the answer option where needed.

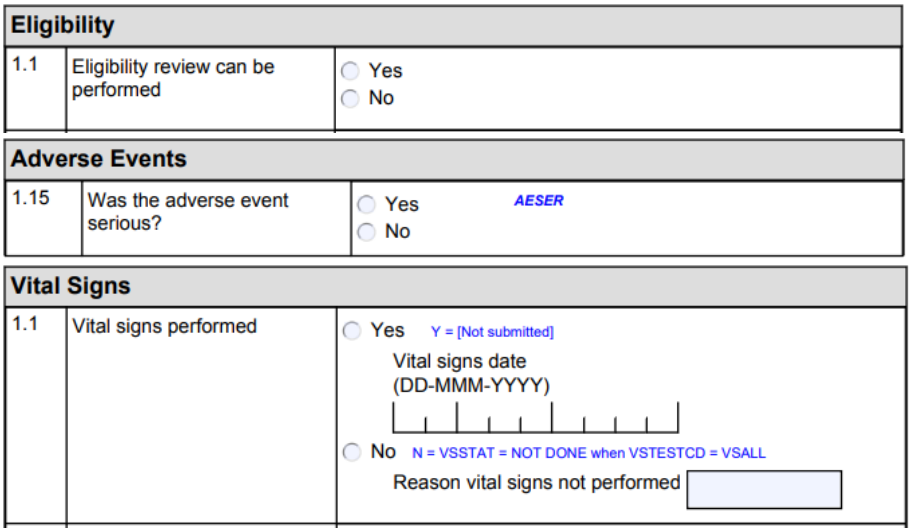

Figure 3. Overview of mapping codes “Yes” and “No” across different forms

STANDARDS REVIEW PROCESS

The standards are built by the data standards team, which is a team within the data management department. This team builds aCRFs, metadata specs and Define-XMLs, is responsible for the review process, and for communication and sharing all required documents with external parties. Once every few weeks, the data standards team shares a batch of forms with the data management leads (DML). They review the forms with the use of an Excel tracking sheet, after which the data standards team processes all feedback. Actions taken based on the DML feedback can be: 1) implement, 2) leave the item as is with an explanation why it was decided to do so, or 3) conclude that the topic needs to be discussed with other disciplines and functional areas.

To discuss these remaining topics, a standards governance team is put in place (figure 4). This team consists of core members and ad hoc team members based on the topics that need to be discussed. With this team, monthly meetings are held to discuss remaining items that need input from other disciplines. All decisions taken during this meeting are documented together with the rationale in order to prevent repeated discussions in the future.

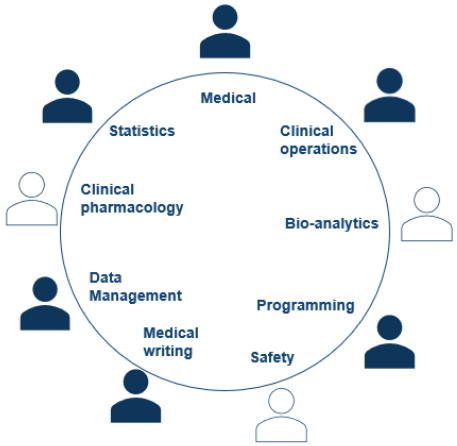

Figure 4. Data standards governance team with representatives from each functional area. Core members are represented in blue, ad hoc members are represented in white.

WHERE ARE WE NOW?

FIRST PACKAGE RELEASED

In September 2022 the first package of aCRFs was released both internally within the company and externally to the data management vendors. One of the most important reasons of building the CRFs from scratch was to carefully reconsider the purpose and use of each item and to create CRFs that are lean and easy to understand. One example that stands out on how taking a step back, and re-evaluating items helps us in making more lean CRFs is the physical examination form where the number of data points collected is nearly halved.

Since we work with different vendors, that have a different organisational structure, we also received different types of feedback on the CRFs released until now. One vendor works with clinical data teams, where data managers and SDTM programmers work together in a team and most of the data cleaning is done in the back-end database, when data is already collected. This is opposite to another vendor, where the data managers and SDTM programmers work in separate teams, and most of the data cleaning is done in the front-end, during data entry. These different approaches come with different feedback and ways of adapting to the new standards.

FROM STANDARDS TO STUDY BUILD

We aim to follow the following process of building study aCRFs and metadata. First, the clinical data manager collects at least the protocol, the version of the SDTM Implementation Guide, and the version of NCI Controlled Terminology to be used for the study.

Based on this, the data standards team starts building the study-specific visit schedule and attaches the required forms to the visits. Together, the clinical data manager and data standards manager identify which forms can be directly imported from the standard and which forms need modifications to adhere to the protocol. The clinical data manager is then responsible for bringing back the CRF casebook to the clinical study team to see if the current design matches the expectations. Once the protocol is finalised and the CRF is updated accordingly, the following documents are planned to be shared with the vendor: the aCRF, the metadata specifications, and the Define-XML. This way of working should support us in creating the CRF in parallel with the protocol in order to align expectations at study level.

CURRENT STATUS

Currently, the first test study build has taken place to identify what part of the process runs smoothly and what still needs to be addressed. The activities within the system of creating visits, setting up a visit schedule and attaching the corresponding CRF pages to the right visits turned out to be a rather easy, and user-friendly process. This exercise reminded us to carefully consider best practices during versioning and assigning identifiers to our CRF building blocks.

CONCLUSION

The implementation of an MDR has taught us a lot. First, it allowed us to create a clear overview of data collected so far. This overview has showed us that we thought we were very consistent, but in some cases it turned out we were not. Second, it has helped us to simplify data collection. Going against the trend of collecting more and more data, we wanted to focus on the work to be done by the sites and simplify the collection. Third, we aim to have a better control of study set-up by standardisation. By doing this we hope to avoid discussions on a critical path. Lastly, we aim to improve the overall quality of our data by having thorough discussions on the data we collect and its purpose.

ACKNOWLEDGMENTS

First, we would like to thank our client for the opportunity to present about our journey of implementing the MDR and all learnings from the past 1,5 year. We would also like to acknowledge our colleagues that are involved in the reviewing rounds. Thank you all for sharing your useful feedback, knowledge and the fruitful discussions. Furthermore, we would like to thank the vendors for helping us to take our standards governance to a next level. And especially to our MDR vendor; thank you for patiently answering all our questions. Last, but certainly not least, we would like to acknowledge our colleagues Isabelle Vandriessche and Jules van der Zalm for their help and useful feedback during the development of this paper and the presentation that goes along with it.

CONTACT INFORMATION

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

Louella Schoemacher
OCS Life Sciences
Ruwekampweg 2G
‘s-Hertogenbosch / 5222 AT
+31 (0)73 523 6000
info.nl@ocs-consulting.com
https://ocs-consulting.nl/ocs-life-sciences/

Jasmine Kestemont
Innovion
Zoetewei 118
Putte / 2850
+32 476 545 917
jasmine.kestemont@innovion.be
www.innovion.be