PhUSE 2017

Paper DS09

Clinical study oversight: different approaches to using data standards

Jasmine Kestemont, Innovion, Belgium
Peter Van Reusel, Innovion, Belgium

ABSTRACT

While most if not all large pharma's have implemented CDISC standards by now, we do see that small to medium size companies and academia have low adoption rates. This may be due to lack of awareness or understanding, but often they simply struggle with an implementation strategy. So how does a company that is new to standards see the wood through the trees in the new regulatory environment? We want to discuss a number of approaches in how to establish data standards between sponsors and CRO's, each with pro's and con's.

INTRODUCTION

Though we can not imagine a daily life without standards, the pharmaceutical world has been a very slow adopter. Standards are still very much seen as something that FDA or PMDA want, so why invest in it when the probability of success for approval for a drug in phase I is less than 10%[1].
Moreover, standards sound complex, and one should leave it to the experts, right?
In the below article, we will describe some benefits to having standards early (or rather, the risks of not having standards), and show some approaches to implementing standards that are not overly complex and will not burn the clinical development budget of even a small biotech company with a very limited portfolio.

APPROACHES TO IMPLEMENTING A LIBRARY

While undoubtedly there are many more models and variants of models than mentioned below, in this article, we will consider models that can work for small to midsize biotechs, pharmaceutical and device companies and academia. There is an assumption that the company either has in-house data standards expertise or has access to expertise through a third party (CRO, technology partner or consultant). The approaches presented below, gradually increase in complexity and associated cost, but also with increasing benefits when implemented well. The article assumes that CRF build and data management work is outsourced to a CRO, but each of the models can work with an in-house team as well.
.

APPROACH 1: THE CRO IS THE EXPERT, LET THEM DECIDE

Most small biotech companies, with a successful "first in human" trial need to start thinking about their Compound Development Strategy, and probably their company growth strategy. At a time when costs are high and Return on Investment is still a distant goal, a company may find it wise to hire a biostatistician, chances are they will not prioritize on hiring a statistical programmer, data manager, let alone a data standards steward.

Nevertheless, many have heard that if they want to submit their data to FDA or PMDA, they will have to adhere to CDISC standards, …. whatever that may be …
So, when writing the RFP for the next trial, CDISC deliverables are required.
All CROs confirm that their deliverables are CDISC compliant, so why does everyone say that standards are complicated?

We will use an example to illustrate the problem:

One of the questions in the protocol's Schedule of Assessments was: "Medical History including history of substance abuse".

The selected CRO handled this question as follows:

HISTORY OF ALCOHOL / DRUG ABUSE			
1	Has the subject every used alcohol? SUTRT = ALCOHOL	SUOCCUR = 'Y' if "Current" or "Former" SUOCCUR = 'N' if "Never"	○ Never ○ Current ○ Former
2	Start Date of Alcohol Consumption	SUSTDTC	__/___/____
3	End Date of Alcohol Consumption	SUENDTC	__/___/____
4	Units of Alcohol Consumed	SUDOSE	___ ○ Per Day ○ Per Week ○ Per Month ○ Per Year SUDOSFRQ

Figure 1: CRF extract of Substance Use

In the meantime, the biotech company had decided to partner with another CRO that had experience with a specific device needed to deliver the drug.
Again, the question on history of substance abuse was raised in the protocol.
Much to the surprise of the biotech, this CRO came up with something very different.

MEDICAL HISTORY [MH]				
BODY SYSTEM		Start Date	End Date	Instructions
1 Dermatological	----------	--/---/----	--/---/----	
2 Respiratory	----------	--/---/----	--/---/----	
3 Psychological	----------	--/---/----	--/---/----	Please include any details on substance abuse
4 HEENT	----------	--/---/----	--/---/----	
5 Cardiovascular	----------	--/---/----	--/---/----	
MHBODSYS	MHTERM	MHSTDTC	MHENDTC	

Figure 2: CRF Extract of Medical History

So which CRO got it right and which didn't?
One can argue that one CRO may have chosen a better approach than the other, and certainly one version will provide a lot more detail than the other, but neither was wrong.

Could this have been avoided? Very likely.
If the Biotech company had created a very simple set of standard forms, which included a Medical History and a Substance Use form, then chances would significantly increase that both CROs would have selected the same form to capture this information.

Is this a problem? Probably not in the case of the history of a subject participating in a trial. It is not very likely that the data would be pooled across trials, unless the drug is intended to treat substance abuse and the population needs to be described.

Is it enough? No.

In the below example, a set of forms was created, and the CROs were asked to use these when building their CRF and SDTM database.

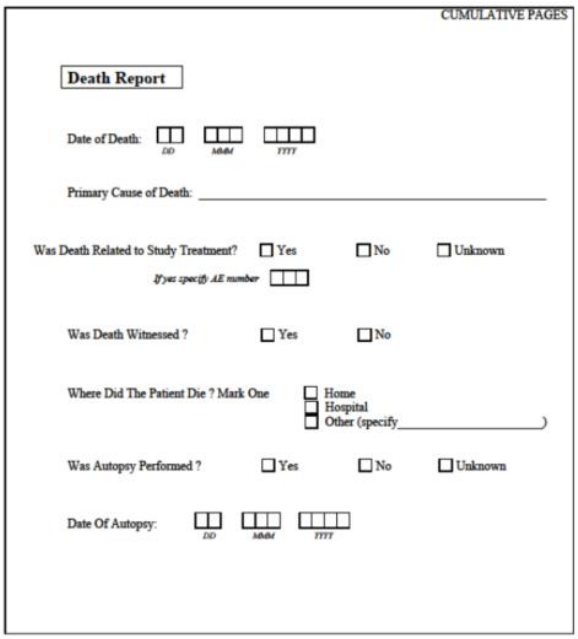


Figure 3: CRF of a Death Report

When the annotated CRF was sent to the Biotech company for review, they looked very different.

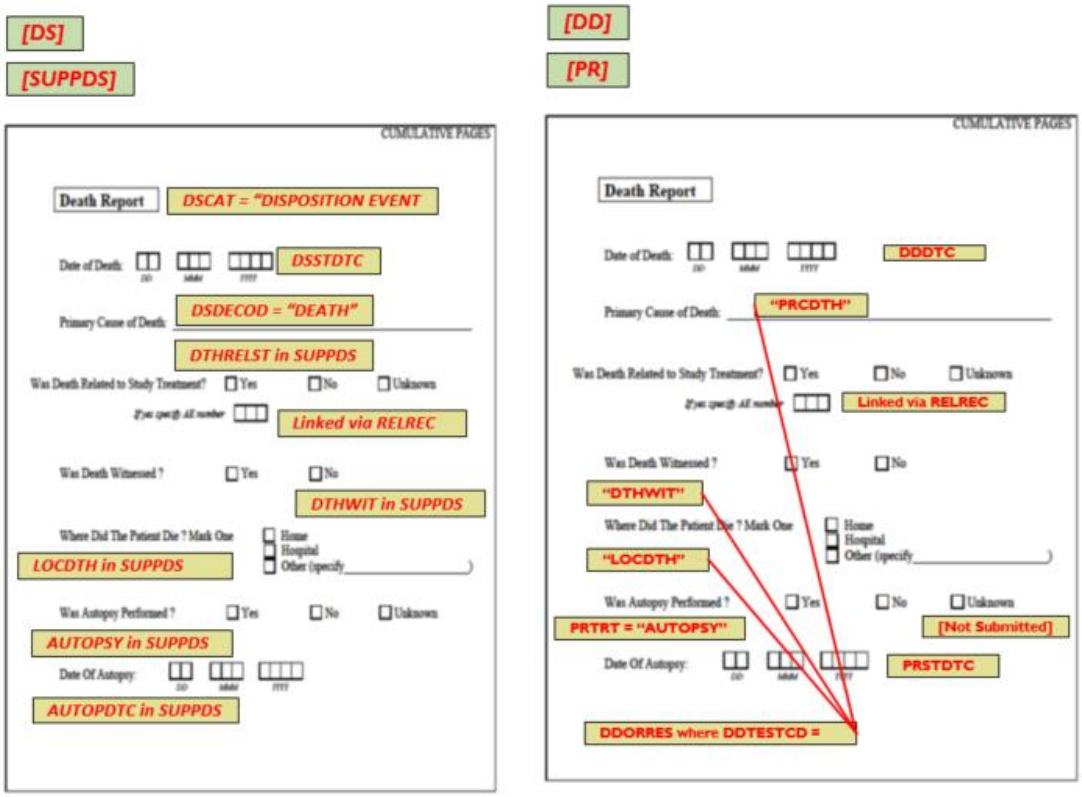


Figure 4: 2 variants of an SDTM annotated CRF of a Death Report

If there had been deaths in both trials, it would have been very challenging to pool the data. The example illustrates that, for two parties (Sponsor and CRO) to "speak the same language", they need to agree on what is captured (the CRF, or in CDISC terminology - CDASH) and how it is represented (SDTM).

While the "light" approach, where the Biotech company relies on the CRO to define the expectations, presents its challenges, there are situations where this model can work. If a Biotech company is committed to a single CRO, where this CRO is really a biometrics extension of the Biotech company, then this is a cheap and effective solution. Most likely the CRO will have invested in their own set of standards and the Biotech company profits from this investment.

APPROACH 2: THE SPONSOR OWNS A BASIC LIBRARY

At a minimum, a "simple" library should contain a CRF layout or other form of data visualization and associated SDTM annotations. This will ensure consistency of data collection and data reporting.

In this case the Sponsor will communicate with the CRO through a study set up package. This package contains a protocol, a (sub)set of CRFs and preferably data specifications.

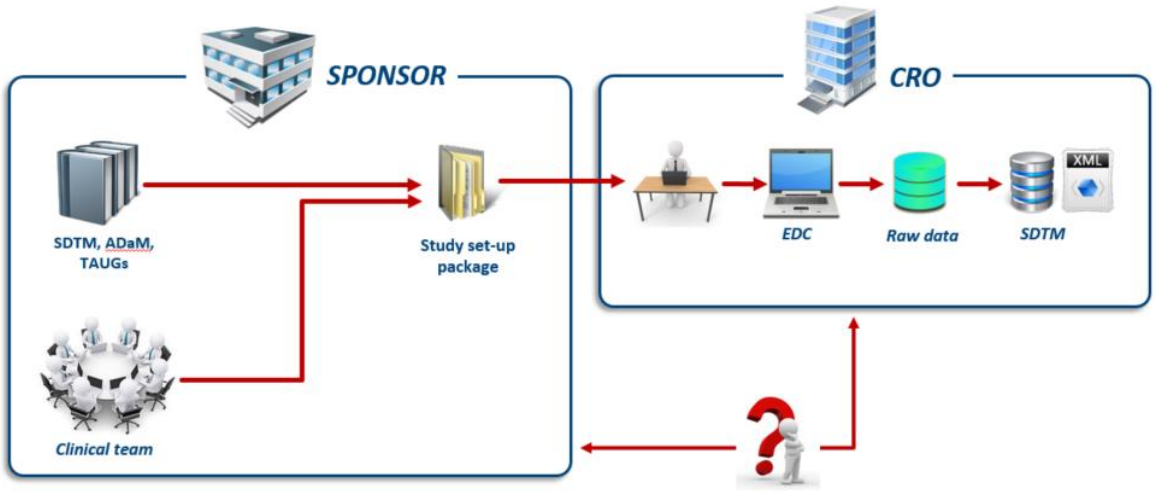
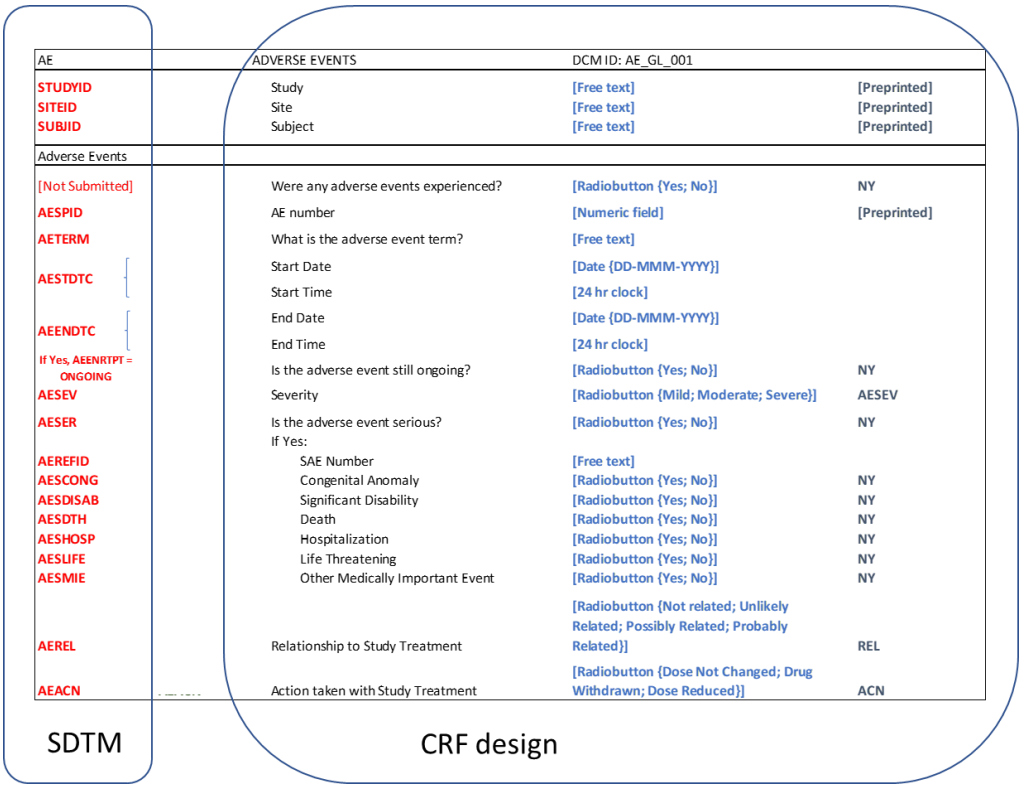


Figure 5: Possible collaboration process between a Sponsor and CRO with a basic library

The data specifications are particularly useful when there are multiple ways of collecting data, and the protocol does not provide the level of detail that a CRO may need (refer to the example of the medical history of substance abuse). A data specification package may be as simple as a list of the unique name and version of CRFs to be used.

How confident can you be that the CRO delivers what you expect?
The advantage of standardization is that there are tools on the market, some of them free of charge, that facilitate the review of your SDTM data structure. While they will not check the content of the data, these tools are instrumental to checking your data for compliance, and will give you some level of confidence that the data is conformant to FDA or PMDA expectations.

When would this be a good model? It is a sustainable option for a small to midsize company, that has a focus on quality and consistency, may plan to bring a portfolio into full development (phase II/III), but has a limited budget for infrastructure. In other words, a library would be a standalone tool, in its most elementary form a versioned word document or excel spreadsheet, as shown in the picture below.



AE	ADVERSE EVENTS	DCM ID: AE_GL_001	
STUDYID	Study	[Free text]	[Preprinted]
SITEID	Site	[Free text]	[Preprinted]
SUBJID	Subject	[Free text]	[Preprinted]
Adverse Events			
[Not Submitted]	Were any adverse events experienced?	[Radiobutton {Yes; No}]	NY
AESPID	AE number	[Numeric field]	[Preprinted]
AETERM	What is the adverse event term?	[Free text]	
AESTDTC	Start Date	[Date {DD-MMM-YYYY}]	
	Start Time	[24 hr clock]	
AEENDTC	End Date	[Date {DD-MMM-YYYY}]	
	End Time	[24 hr clock]	
If Yes, AEENRTPT = ONGOING	Is the adverse event still ongoing?	[Radiobutton {Yes; No}]	NY
AESEV	Severity	[Radiobutton {Mild; Moderate; Severe}]	AESEV
AESER	Is the adverse event serious?	[Radiobutton {Yes; No}]	NY
	If Yes:		
AEREFID	SAE Number	[Free text]	
AESCONG	Congenital Anomaly	[Radiobutton {Yes; No}]	NY
AESDISAB	Significant Disability	[Radiobutton {Yes; No}]	NY
AESDTH	Death	[Radiobutton {Yes; No}]	NY
AESHOSP	Hospitalization	[Radiobutton {Yes; No}]	NY
AESLIFE	Life Threatening	[Radiobutton {Yes; No}]	NY
AESMIE	Other Medically Important Event	[Radiobutton {Yes; No}]	NY
AEREL	Relationship to Study Treatment	[Radiobutton {Not related; Unlikely Related; Possibly Related; Probably Related}]	REL
AEACN	Action taken with Study Treatment	[Radiobutton {Dose Not Changed; Drug Withdrawn; Dose Reduced}]	ACN

SDTM

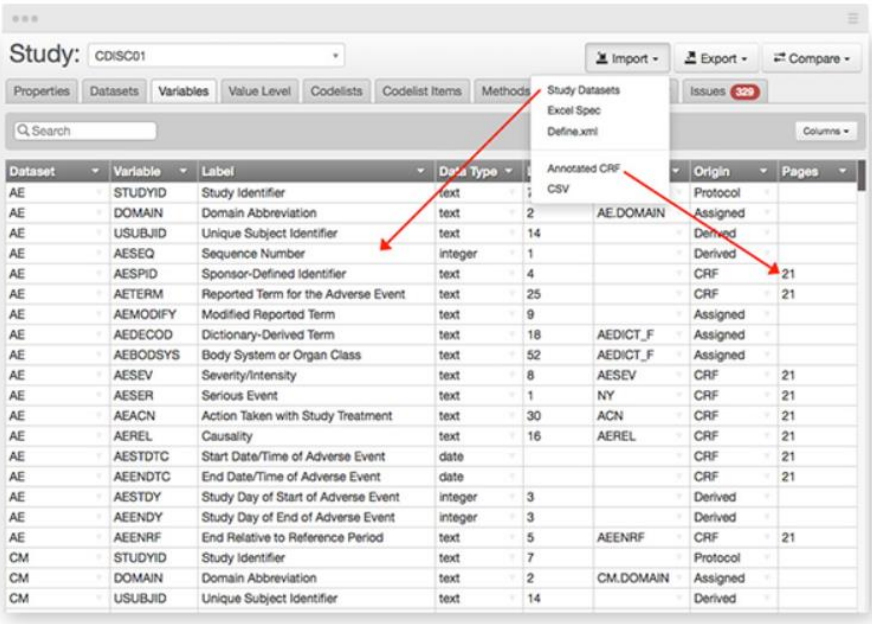
CRF design

Figure 6: Example of a basic library extract

APPROACH 3: THE SPONSOR OWNS AN INTEGRATED LIBRARY

Once standards are implemented, the options for automation, thus building efficiency and quality into your process, are almost endless.

An integrated library not only contains a data visualization and SDTM mapping, but will also have a metadata library. Additional information captured in this library includes codelists, computational methods, value level metadata, etc. Certainly when your library is still small or managed by a limited number of users, it is still possible to maintain the library in Excel. The advantage of using a tool like excel is that the tool is well known (so accessible to anyone) and almost any system can read or upload an Excel file. This means that you have an easily accessible tool, system agnostic tool that gives you the freedom to work with any EDC tool and thus the freedom to work with any CRO partner.



Dataset	Variable	Label	Data Type			Origin	Pages
AE	STUDYID	Study Identifier	text			Protocol	
AE	DOMAIN	Domain Abbreviation	text	2	AE.DOMAIN	Assigned	
AE	USUBJID	Unique Subject Identifier	text	14		Derived	
AE	AESEQ	Sequence Number	integer	1		Derived	
AE	AESPID	Sponsor-Defined Identifier	text	4		CRF	21
AE	AETERM	Reported Term for the Adverse Event	text	25		CRF	21
AE	AEMODIFY	Modified Reported Term	text	9		Assigned	
AE	AEDECOD	Dictionary-Derived Term	text	18	AEDICT_F	Assigned	
AE	AEBODSYS	Body System or Organ Class	text	52	AEDICT_F	Assigned	
AE	AESEV	Severity/Intensity	text	8	AESEV	CRF	21
AE	AESER	Serious Event	text	1	NY	CRF	21
AE	AEACN	Action Taken with Study Treatment	text	30	ACN	CRF	21
AE	AEREL	Causality	text	16	AEREL	CRF	21
AE	AESTDTC	Start Date/Time of Adverse Event	date			CRF	21
AE	AEENDTC	End Date/Time of Adverse Event	date			CRF	21
AE	AESTDY	Study Day of Start of Adverse Event	integer	3		Derived	
AE	AEENDY	Study Day of End of Adverse Event	integer	3		Derived	
AE	AEENRF	End Relative to Reference Period	text	5	AEENRF	CRF	21
CM	STUDYID	Study Identifier	text	7		Protocol	
CM	DOMAIN	Domain Abbreviation	text	2	CM.DOMAIN	Assigned	
CM	USUBJID	Unique Subject Identifier	text	14		Derived	

Figure 7: extract of study metadata from an "integrated library"[2]

Careful design of your metadata library will allow a very simple extraction routine for study builds and generation of define.xml according to the latest standards.

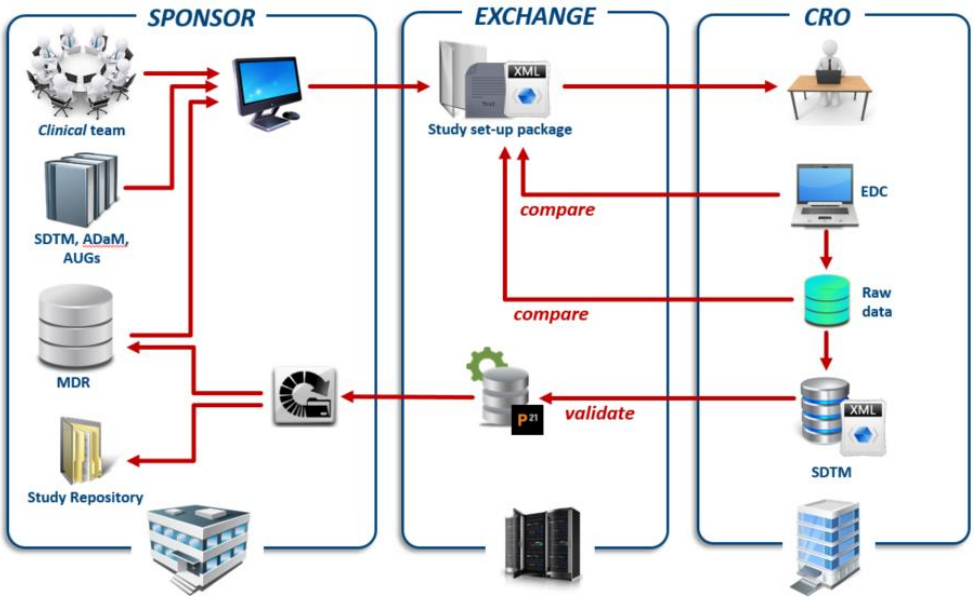


Figure 8: Possible automation process with an integrated library

Once the integrated Library is in place, you can start implementing automations, including, but not limited to automated EDC build and automated validation of your data structure.

CONCLUSION

It is clear that clinical data standards are here to stay, and the sooner in your process you build in the standardization, the greater the benefit, sponsors will still have to make decisions on level of standardization. This will greatly depend on the company's operational model, staff and budget. There is no one size fits all solution, but there is a solution for everyone. Key for success is to look at the long-term goals and processes and implement stepwise. Even a minimal investment can make a huge difference in how CROs and Sponsors communicate and manage expectations.

REFERENCES

[1] https://www.bio.org/sites/default/files/ClinicalDevelopmentSuccessRates2006-2015BIO,Biomedtracker,Amplion2016.pdf
[2] Pinnacle21.com/products/define-xml

CONTACT INFORMATION

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

Jasmine Kestemont
Innovion
Zoetewei 118
2580 Putte
Belgium
+32 473 94 69 92
Jasmine.kestemont@innovion.be
www.innovion.be