Paper RG04

An FDA Submission Experience Using the CDISC Standards

Angelo Tinazzi, Cytel Inc., Geneva, Switzerland
Cedric Marchand, Cytel Inc., Geneva, Switzerland

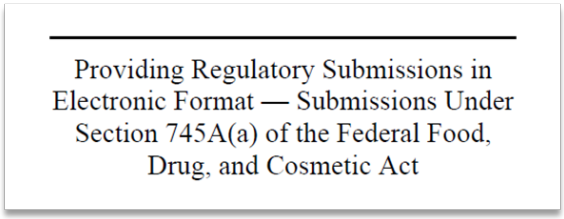
ABSTRACT

The purpose of this presentation is to share an FDA submission experience using the CDISC standards. After introducing the key current requirements when submitting data sets to the FDA, either SDTM or ADaM, some key learning will be shared. This includes, for example, interaction with the FDA and the additional requests we received as well as the feedback after performing the test submission.

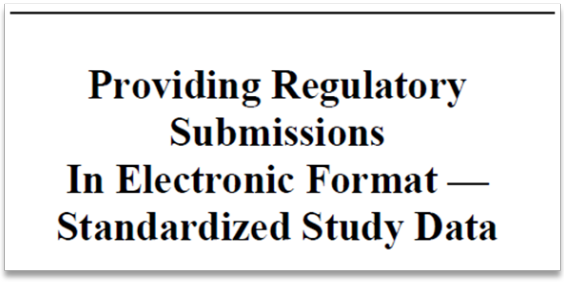
INTRODUCTION

The content of this paper represents our personal experience with this particular submission with this specific sponsor on a specific indication. Although the paper contains information coming from existing requirements, such as CDISC standards and FDA guidance, they represent our experience of applying standards and interacting with the FDA reviewer. Topic and timing of submission, as well as reviewer 'preference', are important factors to consider when submitting data to FDA.

KEY REQUIREMENTS



The parent guidance in this series of documents is the "Guidance for Industry: Providing Regulatory Submissions in Electronic Format – Submissions Under Section 745A(a) of the Federal Food, Drug and Cosmetic Act" [1]. The primary objective of this guidance is to affirm that, as soon as December 2016, you will need to submit most if not all INDs, NDAs, ANDAs and BLAs electronically as opposed to filing on paper.



The second guidance is "Guidance for Industry: Providing Regulatory Submissions in Electronic Format – Standardized Study Data" [2]. Following on to the requirement that most if not all submissions must be electronic, this guidance states that studies initiated in the relatively near future must utilize specific data standards for the collection, analysis and delivery of clinical and non-clinical trial data and results as endorsed by the FDA as documented in the Data Standards Catalog [3]. This requirement kicks in for studies that would support an NDA, ANDA or BLA on the 2 year anniversary the guidance document becoming final (December 17, 2016) and one year later for INDs.

The Study data Technical Conformance Guide [4] provides specifications, recommendations, general considerations on how to submit standardized study data using FDA-supported data standards located in the FDA Data Standards Catalog.

"Sponsors whose studies start after **December 17, 2016**, must submit data in the data formats supported by FDA and listed in the FDA Data Standards Catalog. This applies to NDAs, BLAs, ANDAs, and subsequent submissions to these types of applications."

- FDA U.S. Food and Drug Administration

HOW?

In addition to standard requirements covered by the different CDISC Implementation Guidance, most of the technical requirements are covered by the FDA Study Data Technical Conformance Guide and by the FDA Standards Catalogue where current accepted standards by FDA are listed. The catalog for example lists not only the current CDSIC versions validated and therefore accepted by the FDA, such as SDTM, ADaM and standards controlled terminology, but also the exchange formats to be used such as SAS XPT, XML, PDF, and ASCII, and the additional standard dictionary requirements such as for Adverse Events (i.e. MedDRA).

Furthermore other guidance from CDISC, such as the "CDISC Metadata Submission Guidelines" [7] where for example some recommendations are given for annotating the SDTM aCRF, or the FDA Portable [8] document where detailed requirements are provided for PDF file such as PDF file properties i.e. appearance of bookmarks or file properties. Last but not least the Electronic Common Technical Document (eCTD) [5] contains other details to be considered when naming and organizing files in a specific structure i.e. for file name maximum 64 characters and use only lowercase letters, digits and '-' (hyphen).

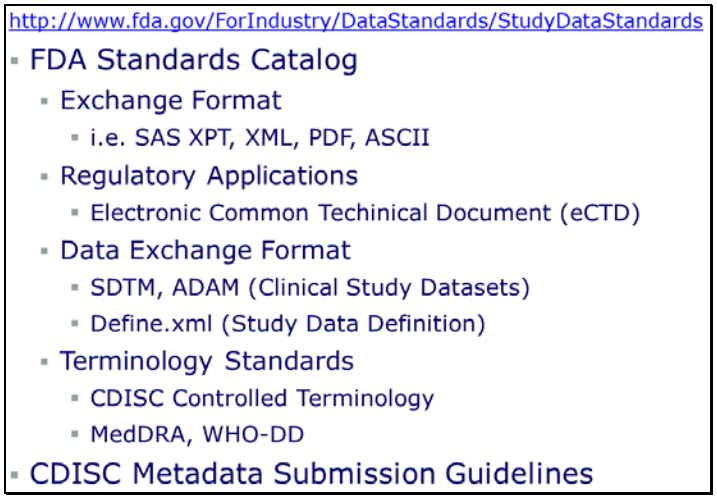


Figure 1: Main standards to be used when submitting data to the FDA

The FDA Technical Rejection Criteria [6] should be also considered when submitted data to FDA, although to date only few are related to datasets:

- for SDTM Trial Summary (TS) and Demographics (DM) dataset are mandatory
- for ADaM, ADSL is mandatory

The TS dataset is also required when non-SDTM datasets are submitted (i.e. legacy datasets).

THE SUBMISSION DATA PACKAGE

As previously mentioned the submission data package should follow a specific folders and files organization [4] [5].

For the clinical part a specific folder is dedicated: the 'm5' folder. Figure 2 shows how our data submission was structured with one folder per study, plus two additional folders containing the ISS and ISE specific files. Within each of these folders the same structure is repeated as shown in figure 2.

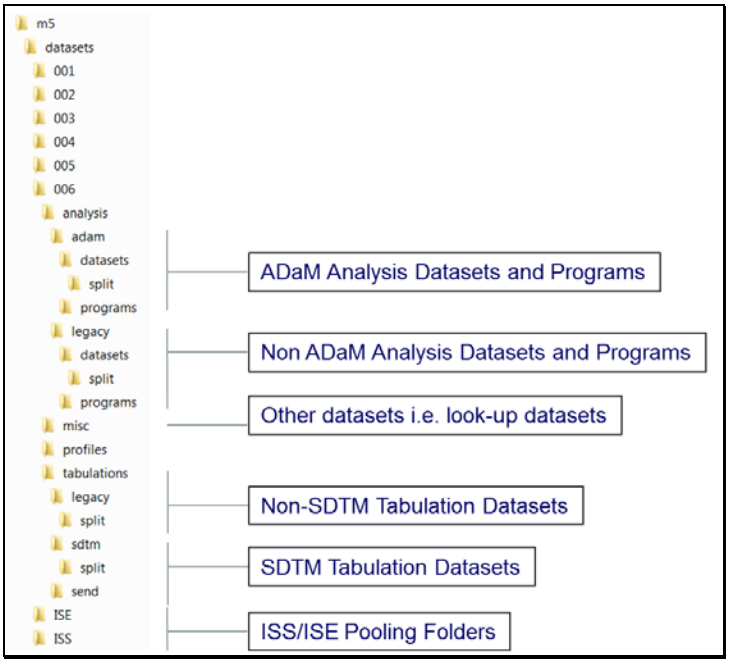


Figure 2: the eCTD m5 folder structure

The data submission package is made of different type of files, such as SAS datasets (xpt files), study data definition (xml files), PDF files and eventually but not required xls files containing the validation reports from for example Pinnacle 21 (see figure 3). Figure 4 shows an example of possible composition of a study folder and ISS/ISE folders where in our case only pooled ADaM datasets were submitted.

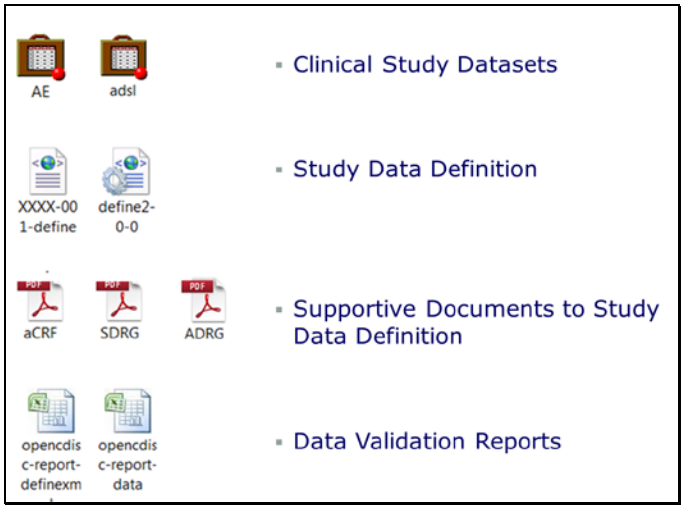


Figure 3: Type of files submitted in the data package

Software programs were also part of the submission (see figure 5). According to the FDA Technical Conformance Guidance we submitted all software programs used to create all ADaM datasets; as for output programs, mainly tables and figures, we submitted all SAS programs. The main purpose of the submission of these programs is to give the reviewer the opportunity to better understand derivations or statistical models used if not enough clear in the documentation provided (i.e. define.xml); as mentioned in the FDA Technical Conformance Guidance *"it is not necessary to submit the programs in a format or content that allow the FDA to directly run the program under its given environment"*. Because we did not submit results metadata we provided high level description of the submitted programs in the Analysis Data Reviewer Guide (ADRG).

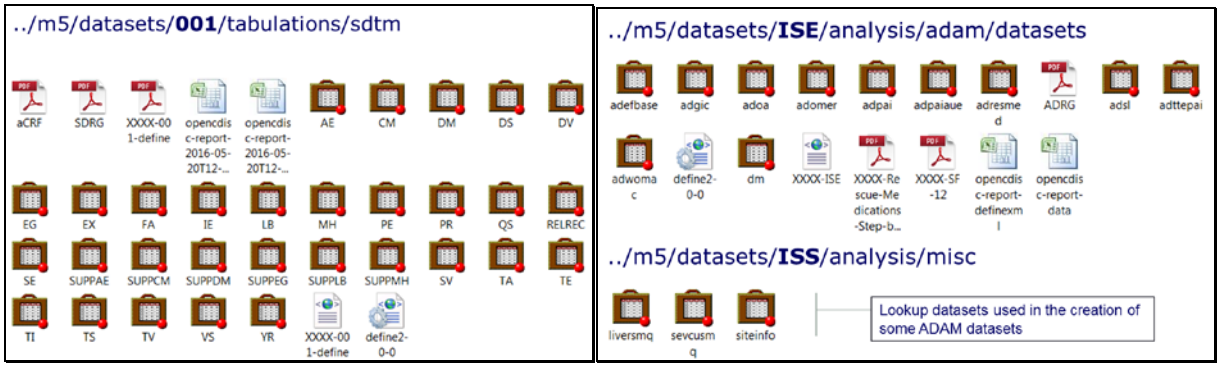


Figure 4: Example of an SDTM study folder and ISS and ISE folder

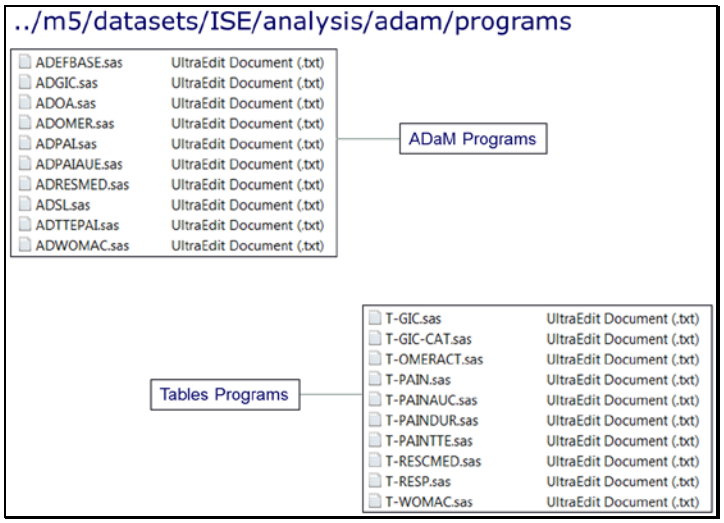


Figure 5: Software Programs

VALIDATION ISSUES

During the validation of the ADaM datasets with Pinnacle21, we came through the issue shown in figure 6. The issue is due to a limitation of FDA Clinical Trial Repository (Janus CTR[1]) system; the database apparently has a maximum length of 1000 characters for data attributes (VARCHAR (1000)). The issue was also discussed in the past in the Pinnacle 21 forum[2]; however apparently in a recent discussion in the LinkedIn group "CDISC-SDTM experts", the issue has been fixed so in the near future the validation checks will be updated.

Domain	Record	Count	Variables	Values	Pinnacle 21 ID	Publisher ID	Message	Category	Severity
DEFINE			Value, Attribute	- When only year was available, OADIAGY was computed as year of study drug administration (ADSL.TRTSDT) – year of diagnosis - When only month and year was available, the day was imputed to 1 and OADIAGY was computed as (study drug administration (ADSL.TRTSDT) - date of Osteoarthritis index knee diagnosis (OADIAGDT)+1)/365.25 - Otherwise if the diagnosis date was not partial the	DD0086		Invalid length of attributes in Define.xml	Structure	Error

Figure 6: ADaM validation issue with long comments

Whether or not the limitation has been removed the recommendation when dealing with long description of complex algorithm such as the one in figure 6, is to either use the Analysis Data Reviewer Guide or to make use of additional documents (i.e. PDF) and reference these documents in the define.xml as shown in figure 7.

[1] Janus CTR is the standard FDA infrastructure that support receipt, validation, storage, easy access and analysis of study data (https://www.fda.gov/ForIndustry/DataStandards/StudyDataStandards/ucm155327.htm)

[2] https://www.pinnacle21.com/forum/dd0086-maximum-length-1000-characters-data-attributes

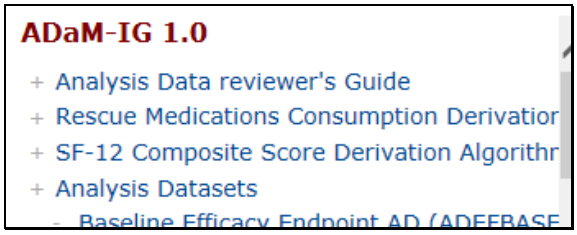


Figure 7: Reference external document in define.xml

OUR RECENT SUBMISSION

The following are the key characteristics of one of our most recent significant submission at the FDA:

- Indication: Pain in a « specific » indication
- Scope of Work: FDA NDA submission
  - ISS: Integrated Summary of Safety
  - ISE: Integrated Summary of Efficacy
- Nr. Of studies: 6
  - 3 only ISE: 1018 Randomized patients
  - 6 ISS: 1155 Randomized patients
  - Screening Failure Patients not included in the SDTM packages
    - → FDA Requested later on 'some' SF data for pivotal studies only [10]

Cytel was involved in the SDTM migration of all submitted studies, the analysis of the Phase II/III pivotal studies, the ISS/ISE pooling and analysis. Moreover, although a specialized company was appointed for the preparation of the entire submission package (eCTD), we provided advices on how to organize the Data Submission package. The sponsor was responsible to Interact with the FDA.

Standards Used

The following standards were used:

- SDTM Ig 3.2
  cSDRG (clinical Study Data Reviewer Guide) as per latest PhUSE template [11]
- ADAM Ig 1.0
  ADRG (Analysis Data Reviewer Guide) as per latest PhUSE template [12]
- Define.xml 2.0 (without results metadata)
- Output programs details were provided in the ADRG i.e. SAS proc used with details on options used (i.e. with PROC MIXED), analysis dataset and selection criteria used for each output (i.e. PARAMCD to be used, way of selecting records to be analysed).

Current Status

Submitted in December 2016, we received the first set of FDA Feedback in February 2017. Since then we entered in a kind of "loop" with FDA asking additional details and new questions, with the sponsor assessing the request and interacting with Cytel to see what other actions are required to properly answer to the FDA reviewer i.e. new exploratory analyses.

Then the good news we received on Friday October 6th, the week just before PhUSE.

INTERACTION WITH THE REVIEWER

Formal meetings between the FDA and Sponsors or Applicants are described in a specific FDA guidance [9].

TYPE OF MEETING

- Type A: a meeting needed to help an otherwise stalled product development program proceed i.e. meetings for discussing clinical holds
- Type B: pre-IND, end-of-Ph-I, pre-NDA
- Type C: any non-type A / Type B meeting regarding the development and review of a product

PRE-NDA MEETING

Figure 8 gives an idea of potential timeline to expect prior to final submission. These are from our experience and they should be not considered standard FDA timeline. At the hypothetical "Month: 0" the sponsor should anticipate Items / questions would like to discuss during the meeting with regards to the application.

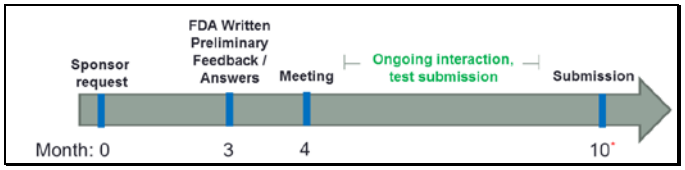


Figure 8: Possible timeline of FDA interaction prior to final submission

The meeting has also the purpose of discussing and find a final agreement on strategy for clinical efficacy and safety in support to the registration, such as type of trials, efficacy endpoints and pooling strategy. It is suggested the sponsor do not use open questions and always propose solutions and ask for confirmation.

Figure 9 shows an example where the sponsor ask the FDA reviewer to confirm they are ok with the submission strategy they intend to follow with regards to the type of studies and data standards they will use and if any legacy datasets with be submitted.

Does the FDA concur with the Sponsor's plan regarding the composition and format of the clinical data submission for the XXXXXXX eCTD NDA?

Table 6. Proposed Specification of Components to be Included in XXXXXX eCTD NDA

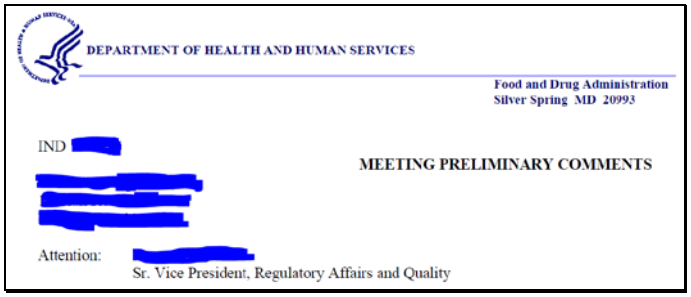
	Tabulations Data and Documentation					Analysis Data and Documentation			
Study	SDTM	Legacy	Define.xml Define.pdf	SDRG	blankcrf .pdf	ADaM	Define.pdf	ADRG	Programs
STUDY 001	×		×	×	×	×	×	×	×
ISS	×		×	×		×	×	×	×
STUDY 002	×		×	×	×				
STUDY 003	×		×	×	×				
STUDY 004	×[a]	×	×	×	×				

Abbreviations: ADaM, analysis data model; ISS, integrated summary of safety; SDRG, Study Data Reviewers' Guide; SDTM, study data tabulation model; TLF, tables, listings and figures

[a] Safety data will be included as part of the integrated ISS database.

Figure 9: Data submission strategy proposed to the FDA by the sponsor

Prior to the meeting usually the reviewer anticipates some feedback and questions that could be discussed during the face to face meeting.


DEPARTMENT OF HEALTH AND HUMAN SERVICES

Food and Drug Administration
Silver Spring MD 20993

IND [redacted]

[redacted]

MEETING PRELIMINARY COMMENTS

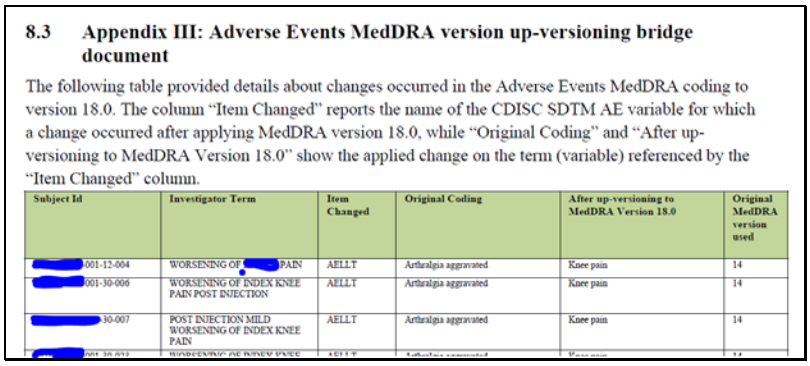
Attention: [redacted]
Sr. Vice President, Regulatory Affairs and Quality

Figure 10: FDA Header Letter

For example in our case, in addition to "punctual" comments to the Statistical Analysis Plan of the ISS and ISE such as suggesting the SMQ to use to further isolate / group the adverse events, they re-iterate the need to have for safety analysis datasets they key demographics and treatment information, and for adverse events information the duration of the adverse event, the outcome, a flag indicating whether or not the event occurred within 30 days of discontinuation of active treatment. This later information was not planned in our analysis datasets and therefore added following the FDA request.

Furthermore for adverse events analysis datasets they specifically asked to include all MedDRA variables such as the lower level term (LLT), the preferred term (PT), the high level term (HLT), etc., including the code for each lower level term. In most of the cases the requirements were already part of the Technical Conformance Guide. One example is the issue of different MedDRA versions in the different studies and the need to have a single version of MedDRA in the pooled ISS analysis datasets. As requested by FDA in their letter we provided a report in each single study SDRG containing the preferred term or the hierarchy mapping changed when the data was converted from one MedDRA version to another. This, as requested by the FDA reviewer, was useful for "understanding discrepancies

that may appear when comparing individual study reports/data with the ISS study report/data" (see figure 11).



8.3 Appendix III: Adverse Events MedDRA version up-versioning bridge document

The following table provided details about changes occurred in the Adverse Events MedDRA coding to version 18.0. The column "Item Changed" reports the name of the CDISC SDTM AE variable for which a change occurred after applying MedDRA version 18.0, while "Original Coding" and "After up-versioning to MedDRA Version 18.0" show the applied change on the term (variable) referenced by the "Item Changed" column.

Subject Id	Investigator Term	Item Changed	Original Coding	After up-versioning to MedDRA Version 18.0	Original MedDRA version used
[redacted]001-12-004	WORSENING OF [redacted]PAIN	AELLT	Arthralgia aggravated	Knee pain	14
[redacted]001-30-006	WORSENING OF INDEX KNEE PAIN POST INJECTION	AELLT	Arthralgia aggravated	Knee pain	14
[redacted]30-007	POST INJECTION MILD WORSENING OF INDEX KNEE PAIN	AELLT	Arthralgia aggravated	Knee pain	14
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Figure 11: Preferred term or hierarchy mapping changes reported in the SDRG

By-Site investigator listings for investigator on-site inspections

FDA uses onsite inspections to ensure that clinical investigators, sponsors, and Institutional Review Boards (IRB) comply with FDA regulations while developing investigational drugs or biologics. Medical reviewers, who are responsible for approving or disapproving a product, consult with BIMO reviewers to choose which clinical trial sites to inspect. For this purpose they requested to provide by-site investigator listings for the two pivotal studies to be used by the FDA Office of Scientific Investigations (OSI) for inspection visits at the selected investigator site [13][14]. See also more details in the appendix.

Additional information/details requested

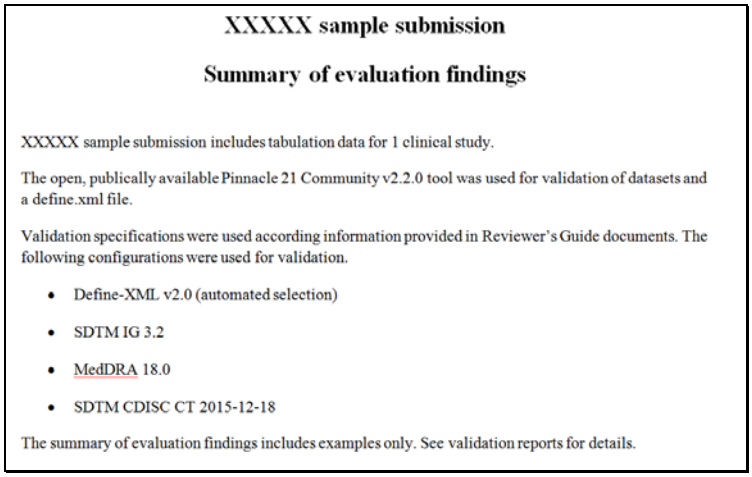
The following is a list of additional information requested by the reviewer:

- laboratory data with normal ranges;
- use of WHO drug dictionary;
- unique coding / nomenclature for Placebo across studies;
- replication of potential covariates / subgroup variables in all ADaM datasets i.e. RACE, SEX. Make a clear plan in the SAP;
- case summaries and CRF for all SAEs, deaths and Discontinuation due to Adverse Events;
- site Level Dataset (optional for now);
- for pivotal studies:
    - number of subjects screened for each site
    - number of subjects randomized for each site, if appropriate
    - number of subjects treated who prematurely discontinued for each site

The appendix contains full details contained in the FDA letter with regards to data being submitted by the sponsor.

TEST (MOCK) SUBMISSION

One study with SDTM and ADaM package was sent by the sponsor as part of the eCTD mock submission.



XXXXX sample submission

Summary of evaluation findings

XXXXX sample submission includes tabulation data for 1 clinical study.

The open, publically available Pinnacle 21 Community v2.2.0 tool was used for validation of datasets and a define.xml file.

Validation specifications were used according information provided in Reviewer's Guide documents. The following configurations were used for validation.

- Define-XML v2.0 (automated selection)
- SDTM IG 3.2
- MedDRA 18.0
- SDTM CDISC CT 2015-12-18

The summary of evaluation findings includes examples only. See validation reports for details.

Figure 12: eData feedback on test submission

A word document with outcome of the submitted test datasets was provided to the sponsor (see figure 12). From the report we understood FDA runs the Pinnacle21 Community tool and at this stage they made use of the SDRG 'only' to check for standards used i.e. SDTM Ig Version, but they did not look at any other more specific detail. However they provided some good and detailed technical feedback (suggestions); for example for the define.xml when origins for all Value Level Metadata (VLM) items within one variable are not the same an Origin for Variable should have a missing value with all details provided on VLM » i.e. when a supplemental qualifier dataset has different information of different type i.e. numeric, text or date.

Furthermore they suggested an alternative way of handling 'Other, specify' race in DM dataset. For example:

« CAMBODIAN » should be represented as « ASIAN »
« NATIVE CANADIAN » should be represented as « AMERICAN INDIAN OR ALASKA NATIVE »
« MIDDLE EAST » and « PALESTINIAN » should be represented as « WHITE »

The SDTM Ig provide different options on how to handle the "Other, specify" field and it leaves to the sponsor the decision on which option to use. However this seems a recurrent request and preferred FDA option. The suggestion here is to map race to DM.RACE according to the CDISC-CT (i.e. by checking synonyms mentioned in the CDISC-CT document) and keep the original race in SUPPDM.

MORE DETAILS ON OUR SUBMISSION

SDTM MIGRATION

SDTM migration could be accomplished by following a rigorous process; this process can be divided into at least 5 main steps:

- Gap analysis
- Understanding source datasets
- Modelling the Migration
- Migration
- Finalize, Validate and Document

The above critical points in migrating legacy data to SDTM have been covered by several presentations [15].
However we want here again to emphasize the importance of the gap analysis prior to start the migration.

Gap Analysis

This is probably the most important step for a successful migration and it has to be completed prior to commencing any migration activity. Having a proper gap analysis does not only give an idea on how complex will be the migration, but most important it gives the possibility to the migration specialist to address well in advance potential issues and, most important, if the specialist is coming from a third party that was not involved in the study development process, it gives the possibility of making an inventory on what is available and what is not. This is extremely important with wider migration with legacy studies conducted by different organizations (CROs), with different conventions applied and sometime in different 'era'. In some circumstances it would be not a big surprise discovering that key documents, such as the most recent CRF, are not available or that key information were not coded in the original source datasets, thus making more complicated the medical coding up-versioning (required for ISS). A Gap analysis should address the following topics and collect the following key information:

- Itemization and evaluation of files to support migration activities
  - Study documents
  - CDISC Standards
  - Company Standards / Company Implementation Guidance
- Validate sample CRF fields versus source data
- Reconcile sample CRFs versus source data
- Comparison of protocol amendments/versions against CRF versions
- External data requirements e.g. central labs
- Clarifies the scope and challenges of migration activities
- Identifies differences in data collection formats

Issues encountered during the SDTM Migration

Harmonization of controlled terminology across study in the submission package
A big effort was needed to try to keep harmonized non-standard terms across studied part of the submission. This is an important step as it will facilitate the integration of the SDTM study datasets into the pooled ADaM package.
An example was the harmonization of the wording for visits as shown in figure 14 or the terminology used for QNAM in the SUPPxx datasets.

Protocol	Phase	Ongoing/ Closed?	Study Subjects	Country	Raw Data	SDTM	Analysis Data	protocol	SAP	aCRF	Blank CRF	CSR
TRIAL 01	Ia	Closed	CF with Pa	Netherlands	present	To migrate	Non-ADaM	v2.0.1 / 2006	v1.1 / 2007	v1.7.4 / 2007	NA	v1.1 / 201301
TRIAL 02	Ib/IIa	Closed	CF with Pa	Hungary, Serbia	Waiting- will get from another vendor	To migrate	Non-ADaM	v1.1 / 2007	v2.0 / 2010	v1.1 / 2007	NA	v1.1 / 201404
TRIAL 03	Ib/IIa	Closed	CF with Pa	Belgium	present	To migrate	Non-ADaM	v1.4 / 2008	v1.2 / 2010	v2.2 / 2008	NA	v1.1 / 201304
TRIAL 04	III	Ongoing	CF with Pa	North America, Europe, Australia, New Zealand	present	v3.1.3	Available	v1.4 / 2009	v.2.1 / 2011	v1.5 / 2010	Available	Not Applicable

Figure 13: Example of items tracked in a Gap Analysis Document

VISIT NUM	VISIT	001	002	005	006	008	009	VISIT (Original as in CRF if different)
1	Screening	√	√	√	√	√	√	Day-21 to -1/SCR in 001
1.1	Day-1		√					Day-14 To -1/Screening in 002
2	Day 1 Baseline	√	√	√	√	√	√	
2.1	Day 2^	√	√				√	
2.2	Day 3		√					
2.3	Day 4		√					
2.4	Day 5		√					
2.5	Week 1	√						Week 1 (Day 8) in 001
2.6	Day 14		√					
2.7	Week 2		√					
2.7	Week 2	√						Week 2 (Day 15) in 001
2.8	Week 3		√					
3	Week 4	√	√		√	√		Week 4 (Day 29) in 001
3.1	Week 5		√					
3.11	Day 42		√					
3.2	Week 6		√	√			√	
4	Week 8	√			√	√		Week 8 (Day 57)

Figure 14: Visits harmonization

Medical History

The original CRFs were containing several diagnosis related information and prior procedures / interventions. For information related to diagnosis we decided to map in MH under a specific category the primary diagnosis and all other diagnosis related information in SUPPMH (this seems the most common approach in the CDISC therapeutic user guidance released so far).

	USUBJID	MHTERM	MHCAT	MHSCAT	MHPRESP	MHOCCUR	MHSTDTC	MHENRTPT
16	[redacted]-008-010-8002	ALLERGY TO ACRYLIC NAILS POWDER	GENERAL MEDICAL HISTORY				2013	ONGOING
17	[redacted]-008-010-8002	[redacted]	PRIMARY DIAGNOSIS	[redacted]N OTHER LOCATIONS				ONGOING
18	[redacted]-008-010-8002	[redacted]	PRIMARY DIAGNOSIS	[redacted] IN OTHER LOCATIONS				ONGOING
19	[redacted]-008-010-8002	[redacted]	PRIMARY DIAGNOSIS		Y	Y	2007	
20	[redacted]-008-010-8007	LAMINECTOMY	GENERAL MEDICAL HISTORY				1972-11	

Figure 15: Primary diagnosis information mapped in MH

Rescue Medications

Rescue medications were collected daily by the patient using an IVRS system and in the CRF when reported to investigator during the hospital visits. The first set of rescue medications were mapped in a sponsor findings domain, while the second were mapped under a specific category in CM.

Adverse Events

In most of the older studies the seriousness criteria were not systematically collected. This is sometime a key information FDA wants to see in the AE dataset. Given the limited number of adverse events occurred in the older studies, it was decided together with the sponsor to extract the seriousness criteria from the sponsor safety database. This might be not a recommend approach with studies such as in oncology, with usually more adverse events, where getting and matching the information from the sponsor safety database together with data coming from the clinical study might be more complicated.

Although this is something that is usually done during the analysis and according to the Statistical Analysis Plan, the Treatment Emergent flag was also derived in SUPPAE dataset as requested by the FDA. The recommendation here is, especially if the SDTM migration is done retrospectively, to try to use the same algorithm in both SDTM and ADaM datasets.

Unscheduled visits and EPOCH derivation

The SDTM Ig gives some suggestion on how to assign VISITNUM to unplanned visits in order to maintain chronology. The PhUSE CSS initiative provides some more detailed approaches for VISTNUM and EPOCH derivation [16]. Be aware deriving VISITNUM for unscheduled visits can be very time-consuming.

Data issues of locked studies

One common situation when migrating legacy data to SDTM retrospectively, is that you might find some data issues that you could not fix (all databases are locked). Again here the SDRG is the right place where you could be transparent and describe any data issues detected during the SDTM migration whether or not detected by the validation.

Although not recommended, for obvious mistake where there was a clear evidence of the mistake, in agreement with the sponsor we hard-coded the data and corrected the obvious mistakes. For example a start date after the end date but clearly wrong year or a confirmation obtained from source document without unlocking data. In this case the recommendation is to create a «Note to File», either in the programming or in the data-management documentation, where the correction and rationale are described. Such Note to file can be then referenced in the SDRG.

POOLING FOR ISS AND ISE

The pooling to support the Integrated Summary of Safety (ISS) and the Integrated Summary of Safety (ISE) was done in ADaM from single study SDTMs.

Different MedDRA Versions

All Adverse Events medical coding were up-versioned to a common and more recent MedDRA version. As previously discussed this is required by the FDA as per the Study Data Technical Conformance Guidance and the requirement was re-iterated in the feedback FDA provided prior to the pre-NDA meeting "*If the version that is to be used for the ISS is different than versions that were used for individual study data or study reports, it is important to provide a table that lists all events whose preferred term or hierarchy mapping changed when the data was converted from one MedDRA version to another. This will be very helpful for understanding discrepancies that may appear when comparing individual study reports/data with the ISS study report/data*". Any discrepancy to the medical coding used in the single study original analysis was listed in the SDRGs i.e. any major change such as change of the preferred term.

Differences with Clinical study reports

It is not uncommon that when defining the analysis approach for the pooled analysis, either ISS or ISE, few changes might occur in the way data are derived and analyzed. For example in the 2 pivotal studies, and therefore in the ISE, the way data were windowed were causing some discrepancies compared to an older study. Such a difference was documented and justified in the ISE ADRG (see figure 16).

ISE ADAM Study CSR Table Nr	Discrepancy and Reason
ADPAI 001 14.2.1.x and 14.2.2.x (all pain endpoint summaries)	The visit / week windowing in 001 final analysis was different i.e. week 1 was from day 1 to day 8 whereas in ISE (see SAP section10.1.1) we applied the windowing from day 1 to day 7 as we did for the analysis of study 006 and 008.

Figure 16: Documenting differences in analysing data

CONCLUSION

From our recent FDA data submission experience using the CDISC standards we would like to make the following recommendations:

- adopt CDISC as soon as possible, starting with CDASH. 'Lost in Traceability' has a cost;
- adopt early in advance a good Vendor Surveillance process to make sure your CROs are doing it right and consistently within your project;
- try to influence your sponsor or your regulatory department (educate) on data submission requirements;
- plan ahead and clarify with the FDA reviewer any doubts / special situations (reviewer preferences). For example how to handle analysis of extension studies when pooling information from the parent study;
- a lot of documentation effort (cSDRG and ADRG).

REFERENCES

Available at the FDA Study Data Standards Resources webpage (https://www.fda.gov/forindustry/datastandards/studydatastandards/default.htm)

[1] FDA Guidance for Industry: Providing Regulatory Submissions in Electronic Format – Submissions Under Section 745A(a) of the Federal Food, Drug and Cosmetic Act

[2] FDA Guidance for Industry: Providing Regulatory Submissions in Electronic Format – Standardized Study Data

[3] FDA Data Standards Catalog – FDA

[4] FDA Data Technical Conformance Guide

[5] Electronic Common Technical Document (eCTD)

[6] FDA Technical Rejection Criteria for Study Data

Other useful papers and guidance

[7] CDISC Metadata Submission Guideline (MSG) for SDTMIG https://www.cdisc.org/standards/foundational/study-data-tabulation-model-implementation-guide-sdtmig/metadata-submission

[8] Portable Documentation Format (PDF) Specification https://www.fda.gov/downloads/Drugs/UCM163565.pdf

[9] Formal Meetings Between the FDA and Sponsors or Applicants https://www.fda.gov/downloads/drugs/guidances/ucm153222.pdf

[10] H.B. Winsor, M. Widel. Good versus better SDTM: Including Screen Failure Data in the Study. PharmaSUG, 2017.

[11] Study Data Reviewer Guide (SDRG), PhUSE http://www.phusewiki.org/wiki/index.php?title=Study_Data_Reviewer%27s_Guide

[12] Analysis Data Reviewer Guide (ADRG), PhUSE http://www.phusewiki.org/wiki/index.php?title=Analysis_Data_Reviewer%27s_Guide

[13] FDA OSI Webinar: Overview of information Requested by CDER OSI https://collaboration.fda.gov/p44198603

[14] T. Dreyer, T. Scetinina, OSI Packages: What you need to know for your next NDA or BLS Submission. PharmaSUG, 2015

[15] A. Tinazzi, Looking for SDTM Migration Specialist. PhUSE, 2014.

[16] PhUSE - Best Practices - Assigning VISITNUM to Unscheduled Visits and Assigning EPOCH to Observations http://www.phusewiki.org/wiki/index.php?title=VISITNUM_and_EPOCH

RECOMMENDED READING

D.C. Izard,Preparing, Legacy Format Data for Submission to the FDA When & Why Must I Do It, What Guidance Should I Follow? PharmaSUG, 2016.

M. Stackhouse, T.J. Peterson, Achieving Clarity through Proper Study Documentation: An Introduction to the Study Data Reviewer's Guide (SDRG). PharmaSUG, 2016.

K. Lee. How will FDA Reject non-CDISC submission? PharmaSUG, 2017.

A. Tinazzi, Lost» in Traceability, from SDTM to ADaM …. finally Analysis Results Metadata. CDISC Europe Interchange, 2016.

S. Minjoe, T. Petrowitsch, Traceability: Plan Ahead for Future Needs. PhUSE, 2014.

Y. Nakajima, T. Kitahara. Awareness from Electronic Data Submission to PMDA and FDA -- Lesson & Learnt from hands-on experiences. PharmaSUG, 2017.

S. Sirichenko, M. Kanevsky. Good Validation Practice. PharmaSUG, 2017.

S. Minjoe. Preparing Analysis Data Model (ADaM) Data Sets and Related Files for FDA Submission. PharmaSUG, 2017.

CONTACT INFORMATION

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

Your comments and questions are valued and encouraged. Contact the author at:
Angelo Tinazzi / Cedric Marchand
Cytel Inc.
Route de Prè-Bois 20
1215 Geneva – Switzerland
Email: angelo.tinazzi@cytel.com / cedric.marchand@cytel.com
Web: www.cytel.com

APPENDIX

PRE-NDA MEETING – DATA SUBMISSION STRATEGY FEEDBACK

Sponsor Question: « *Does the Division agree with* Wonderful Company's *proposed plan for submitting standardized electronic datasets for data from the clinical development program?* »

FDA Answer: « *In addition we have the following comments regarding the dataset:* »

1. **The integrated safety dataset must include the following fields/variables:**
   a. **unique patient identifier,**
   b. **study/protocol number,**
   c. **patient's treatment assignment, demographic characteristics, including gender, chronological age (not date of birth), and race,**
   d. **dosing at time of adverse event,**
   e. **dosing prior to event (if different),**
   f. **duration of event (or start and stop dates),**
   g. **days on study drug at time of event,**
   h. **outcome of event (e.g., ongoing, resolved, led to discontinuation),**
   i. **flag indicating whether or not the event occurred within 30 days of discontinuation of active treatment**
   j. **marker for serious adverse events**
   k. **verbatim term**
2. **The adverse event dataset must include the following MedDRA variables: lower level term (LLT), preferred term (PT), high level term (HLT), high level group term (HLGT), and system organ class (SOC) variables. This dataset must also include the verbatim term taken from the adverse event data set and provide a variable that gives the numeric MedDRA code for each lower level term on the case report form.**
3. **The preferred approach for dealing with the issue of different MedDRA versions is to have one single version for the entire NDA. If this is not an option, then, at a minimum, it is important that a single version of MedDRA is used for the ISS data and ISS analysis. If the version that is to be used for the ISS is different than versions that were used for individual study data or study reports, it is important to provide a table that lists all events whose preferred term or hierarchy mapping changed when the data was converted from one MedDRA version to another. This will be very helpful for understanding discrepancies that may appear when comparing individual study reports/data with the ISS study report/data.**
4. **Provide a detailed description for how verbatim terms were coded to lower level terms according to the ICH MedDRA Term Selection: Points to Consider document. For example, were symptoms coded to syndromes or were individual symptoms coded separately.**

5. **Perform the following SMQ's on the ISS adverse event data and include the results in your ISS report: 1. Severe cutaneous adverse reactions SMQ and 2. Possible drug related hepatic disorders – comprehensive search SMQ. Also, provide any additional SMQ that may be useful based on your assessment of the safety database. Be sure the version of the SMQ that is used corresponds to the same version of MedDRA used for the ISS adverse event data.**

6. **The spelling and capitalization of MedDRA terms must match the way the terms are presented in the MedDRA dictionary. For example, do not provide MedDRA terms in all upper case letters.**

7. **For the concomitant medication dataset, you must use the standard nomenclature and spellings from the WHO Drug dictionary and include the numeric code in addition to the ATC code/decode.**

8. **For the laboratory data, be sure to provide normal ranges, reference ranges, and units as well as a variable that indicates whether the lab result was from the local lab or central lab. Also, the variable for the laboratory result must be in numeric format.**

9. **Perform adverse event rate analyses at all levels of MedDRA hierarchy (except for LLT) and also broken down by serious versus non-serious.**
10. **Across all datasets, the same coding must be used for common variables, e.g. "PBO" for the placebo group. Datasets must not incorporate different designations for the same variable, e.g. "PBO" in one dataset, and "0 mg" or "Placebo," in another datasets. If the coding cannot be reconciled, another column using a common terminology for that variable must be included in the datasets.**

11. **All datasets must contain the following variables/fields (in the same format and coding):**
    a. **Each subject must have one unique ID across the entire NDA**
    b. **Study number**
    c. **Treatment assignment**
    d. **Demographic characteristics (age, race, gender, etc.)**

12. **A comprehensive listing of patients with potentially clinically significant laboratory or vital sign abnormalities must be provided. A listing must be provided of patients reporting adverse events involving abnormalities of laboratory values or vital signs, either in the "investigations" SOC or in an SOC pertaining to the specific abnormality. For example, all AEs coded as "hyperglycemia" (SOC metabolic) and "low blood glucose" (SOC investigations) should be tabulated. The NDA analyses of the frequency of abnormalities across treatment groups is not sufficient without ready identification of the specific patients with such abnormalities. Analyses of laboratory values must include assessments of changes from baseline to worst value, not simply the last value.**

For each pivotal trial: Site-specific individual subject data listings (hereafter referred to as "line listings"). For each site, provide line listings for:

a. Listing for each subject consented/enrolled; for subjects who were not randomized to treatment and/or treated with study therapy, include reason not randomized and/or treated
b. Subject listing for treatment assignment (randomization)
c. Listing of subjects that discontinued from study treatment and subjects that discontinued from the study completely (i.e., withdrew consent) with date and reason discontinued
d. Listing of per protocol subjects/ non-per protocol subjects and reason not per protocol
e. By subject listing of eligibility determination (i.e., inclusion and exclusion criteria)
f. By subject listing, of AEs, SAEs, deaths and dates
g. By subject listing of protocol violations and/or deviations reported in the NDA, including a description of the deviation/violation
h. By subject listing of the primary and secondary endpoint efficacy parameters or events. For derived or calculated endpoints, provide the raw data listings used to generate the derived/calculated endpoint.
i. By subject listing of concomitant medications (as appropriate to the pivotal clinical trials)
j. By subject listing, of testing (e.g., laboratory, ECG) performed for safety monitoring

2. We request that one PDF file be created for each pivotal Phase 2 and Phase 3 study using the following format: