PharmaSUG 2021 - Paper SS-110

Assembling Reviewer-friendly eSubmission Packages

Shefalica Chand, Seagen Inc., Bothell WA

ABSTRACT

One thing 2020 has taught us and reenforced is the importance of getting novel treatments out faster to the patients. Pharma and Biotech companies are constantly racing against time to develop new therapies as fast as possible, but once the drug is submitted for regulatory review, the clock switches hands to FDA for review and decision-making. A fair question, then, is what Biometrics staff preparing submissions could do in advance to help facilitate and expedite this review and decisioning process.

FDA regularly shares guidelines, specifications, recommendations, and general considerations for sponsors to submit standardized study data for animal and human trials in the form of documents like the Study Data Technical Conformance Guide, Electronic Common Technical Document (eCTD), FDA Data Standards Catalog, etc. Apart from these guidance documents and industry practices, there are several methods and good practices sponsors can adopt to prepare an optimally accessible eSub package that is friendlier for FDA reviewers.

This paper will share some of those practical processes that can help expedite the review by improving package navigation and proactively submitting helpful non-standard components, which can in turn help reduce unnecessary FDA questions and resulting delays during the review. We will be sharing practical examples, best practices, list of dos and don'ts, useful checklists, etc., to help teams prepare an optimally "reviewer-friendly" eSub package.

INTRODUCTION

Submission of an electronic submission package (eSub) to FDA is a critical milestone in a drug development lifecycle. There are several resources, guidelines, standards, and tools offered by FDA, PHUSE and Pinnacle 21, that help sponsors prepare a robust submission package. We will look into some of the good practices that may result in creating a positive experience for the reviewers and may facilitating faster decisioning.

Main components that are included in a standard eSub package are as listed below:

- SDTM annotated CRF
- SDTM and ADaM data sets (as SAS® V5 transport files)
- define.xml and stylesheet
- Reviewer's Guides (cSDRG and ADRG)
- Programs in ASCII format
- Any additional definitions or supplemental documents/files

There is scope in each of these components to add efficiencies. With extensive experience sponsors can learn some nifty and practical solutions to create an eSub package that can help facilitate a more efficient review by FDA. This can be done by putting ourselves in the reviewers' shoes.

SDTM ANNOTATED CRF

An SDTM aCRF document provides a clear mapping from raw (usually EDC) data to SDTM domains. There are several guidelines in the industry on how to prepare this document, such as CDISC's SDTM aCRF Guideline that shares technical details.

Here are some additional tips to make a more robust aCRF document:

1. All aCRF variables should be either annotated with SDTM variables/domains or labeled [Not Submitted], for clarity to reviewers.
2. If information is "Not Submitted", it is important to double-check it is not relevant information that should have been submitted, and FDA may come back asking for.
3. If common variables like STUDYID and VISIT are annotated without data set names, there should be a relevant label in the document such as "Applicable to all data sets".
4. The font style, size and color must keep reviewers' ease and readability in mind.
    a. **Standard font style:** Arial, Courier, Times New Roman, etc.
    b. **Font sizes** should be **between 9 - 12 points**.
    c. **Font color** for the text should be **black**.
    d. **Annotation box background color**: The colors may vary from one computer monitor to another; thus, it may be difficult to know if reviewer will see the same color as in the original aCRF. The CMYK (Cyan, Magenta, Yellow, Black) color model provides more control and consistency for printing, compared to the RGB model.
    e. **Lighter or pastel colors** will enhance contrast and readability.
5. It is important to **flatten the aCRF file**, to make it non-editable to avoid accidental changes.

STRUCTURE OF SDTM AND ADAM DATA SETS

The data set submissions in an eSub need to follow the guidelines provided by FDA in the Study Data Technical Conformance Guide (TCG) and CDISC standards. Tools like Pinnacle 21 (P21) help validate the CDISC compliance of the data, as well as promote eSub package quality. Any deficiencies and gaps in the data need to be resolved with ongoing data cleaning efforts and programming updates and checks. Any unresolved issues need to be appropriately documented in respective reviewer's guides.

Appropriate planning at early developmental stages of data set programming can help create data sets that are not just compliant with expected standards but also better support the analyses and review.

VARIABLE NAMES AND VALUES

Here are some tips on variable naming and data set values. These may or may not trigger P21 alert messages, however, following these tips can help proactively create robust data sets.

1. If any non-CDISC variable names and labels (example: in SUPP-- or ADaM data sets) are used, make them relevant and relatable to their meaning and analysis.
2. Reserve variable names ending with “--FN” or “--FL” for flag variables that have values such as 0/1 or Y/N, to make these more intuitive for reviewers.
3. Similarly, reserve variable names ending with “--DT”, “--DTM”, “--DY” for date/time/day variables.
4. Companion/paired variable names should be similar, for example XYZC and XYZN, especially if they are non-standard. This will not be identified in P21 checks, however not doing so may add some confusion and inefficiencies during the review process.
5. There should be a 1-on-1 correlation between paired variables.
6. For variables like --TEST or –PARAM, any sponsor-defined values should be relevant to the analysis.
7. Any non-standard variable labels and values/text such as in --TEST or –PARAM should preferably be in title case.
8. If any abbreviations are needed in variable values or labels, they should be easy to understand. Good practice when shortening or abbreviating is to remove vowels followed by “.” so it is clear the word has been abbreviated. However, abbreviations of some regular words should resort to commonly used abbreviations, for example: Avg., Max., Min., Approx., Tot.
9. ADaM data set variables should be "one proc away" based on the analysis. This will help reviewers use ADaM data sets directly without having to do much post-processing or merging with other data sets.
10. Wherever possible, the variable values should be aligned with CDISC controlled terminology values. Any values outside of CDISC CT can be added as sponsor-defined values in extensible CT.
11. Avoid using underscores in variable names to limit unnecessary characters using up space towards the allowable limit of 8 characters in variable names.

DEFINE.XML

Similar to other submission components discussed in previous sections, define.xml documents should also follow the standards and guidelines provided by FDA in the TCG and CDISC standards, and validated using Pinnacle 21 (Community or Enterprise versions).

Here are some additional pointers and good practices that may help make define.xml more reviewer friendly.

WRITING DERIVATIONS AND SPECIFICATIONS FOR FDA REVIEWERS

The Comment/Derivation section in define.xml is read by FDA reviewers who may or may not have extensive SAS® programming background or knowledge of complex programming syntax. Also, company’s internal jargon can cause confusions for reviewers.

To facilitate an efficient review, this section needs to be written in relatable and simplified English without extensive use of SAS® codes/functions or use of arithmetic operators. The logic of each derivation must be described in such a way that reviewers can reproduce the same results following the derivations. Here are a few guidelines that can be followed for writing clear derivations/specifications:

1. Spell out the data set type or SAS® libnames.
   Examples:
   r → CRF Data Set
   s → SDTM Data Set

a → ADaM Data Set
l → Lookup Table

2. Other study-specific SAS® library references should be described similarly.
3. Use upper case for variable and data set names. This will make it easier for reviewers to identify them, amidst all other information and text.
4. Spell out operators and symbols.
   Examples:
   ≥ → greater than or equal to
   ≤ → less than or equal to
   = → equal to (or equals)
   ≠ → not equal to (or does not equal)
   . or " " → is null or is missing

5. Use double quotes for text values.
6. Avoid using any names of team members or companies such as "As per X <team member> the algorithm was defined as…". That information is not needed by FDA reviewers.
7. Run a spellcheck on the specification file prior to creating define.xml, to catch and correct obvious typos.

TABLE 1: MORE EXAMPLES

Variable	Programming Comment	FDA Comment for define.xml
STUDYID	='Study-X'	Equals "Study-X".
USUBJID	='Study-X-10001-'||r.dose.id, where leading zeros are added to keep four digits for patient number	Combine "Study-X-10001-" with ID variable from CRF DOSE data set, where leading zeros are added to keep four digits for patient number.
TRTSDT	=a.adsl.trtsdt	Equal to TRTSDT variable from ADSL data set. Display in date9. format.
AVISIT	The variable name, r.dose.(c1-c16), corresponds to the visit number.	Equals to the treatment cycle as identified on the dosing CRF (Cycle 1 through Cycle 16).
AVISITN	Converted from AVISIT, keep digits only	Equals the numeric portion of the AVISIT variable
PARAM	='Dose per Administration (mg)' for dose per administration; ='Cumulative Dose' for cumulative dose administered	1. For dose administration records, PARAM equals "Dose per Administration (mg)". 2. For cumulative dose records, PARAM equals "Cumulative Dose".
AVAL	=r.dose.(c1-c16) for treatment dosage. Keep digits only. Calculate the sum across r.dose.(c1-c16) for the cumulative dosage	1. For dose administration records, AVAL equals to variables (C1 to C16) from CRF DOSE data set. 2. For cumulative dose records, AVAL equals to the sum of values of variables C1 through C16 from CRF DOSE data set.
EXDOSU	='mg'	Equal to "mg".
DTYPE	='TOTAL' for cumulative records and null value for rest of the records	Equal to "TOTAL" for cumulative records, otherwise equal null for remaining records.

SPONSOR-DEFINED CONTROLLED TERMINOLOGY

Consider adding sponsor-defined controlled terminology for as many variables as possible, especially variables such as Flag, Visit, Epoch, Timepoint, Indicators, etc. This will help FDA reviewers get better insight into the variable values and help efficient review.

LINKING REVIEWER'S GUIDE AND OTHER SUPPLEMENTAL DOCUMENTS

1. The order of Supplemental documents should be based on the priority of the documents.

   Example: The Reviewer's Guide should be the first document in the list.

2. It will also be helpful to reviewers to mention intent of the supplemental document in the list rather than just the name of the file.

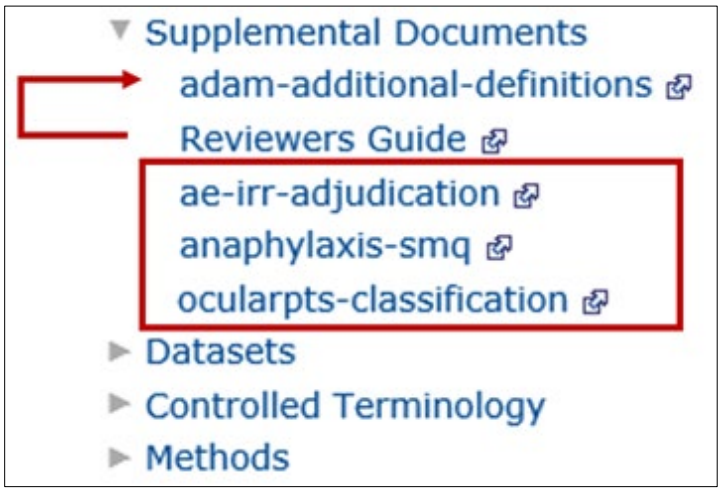

Figure 1: Supplemental Documents 'Before' Updates

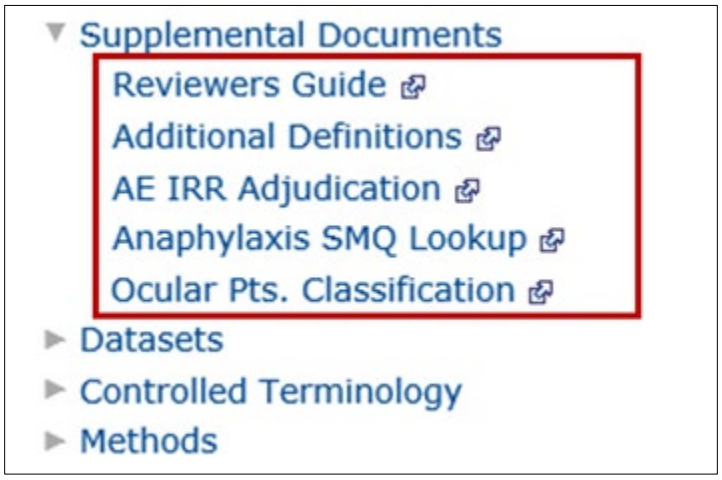

Figure 2: Supplemental Documents 'After' Updates

REVIEWER'S GUIDE

The data reviewer's guide is helpful documentation for regulatory reviewers to understand the background and big picture regarding the study data, and to facilitate a more efficient review, thus it is important to document information in a clear and concise manner with the right amount of detail.

PHUSE offers standard templates for the cSDRG (for SDTM) and ADRG (for ADaM) that are appropriately sectioned, to help summarize all essential details about the data and submission components.

In this section we will share some best practices and tips about creating a more effective reviewer's guide.

1. It is helpful to add the study design screenshot from the study protocol to the reviewer's guide in its protocol design section. This small addition will help reviewers not go back and forth between documents.
2. If a table flows to the multiple pages, repeat the column headers on each page for ease of review.
3. As in define.xml, write any variables or data sets in upper case so they are easily noticeable, otherwise it gets lost in rest of the text.
4. For the "Data Conformance Summary" section, the page orientation should be landscape to match the PHUSE template. It improves the readability of the content, especially issue explanations.

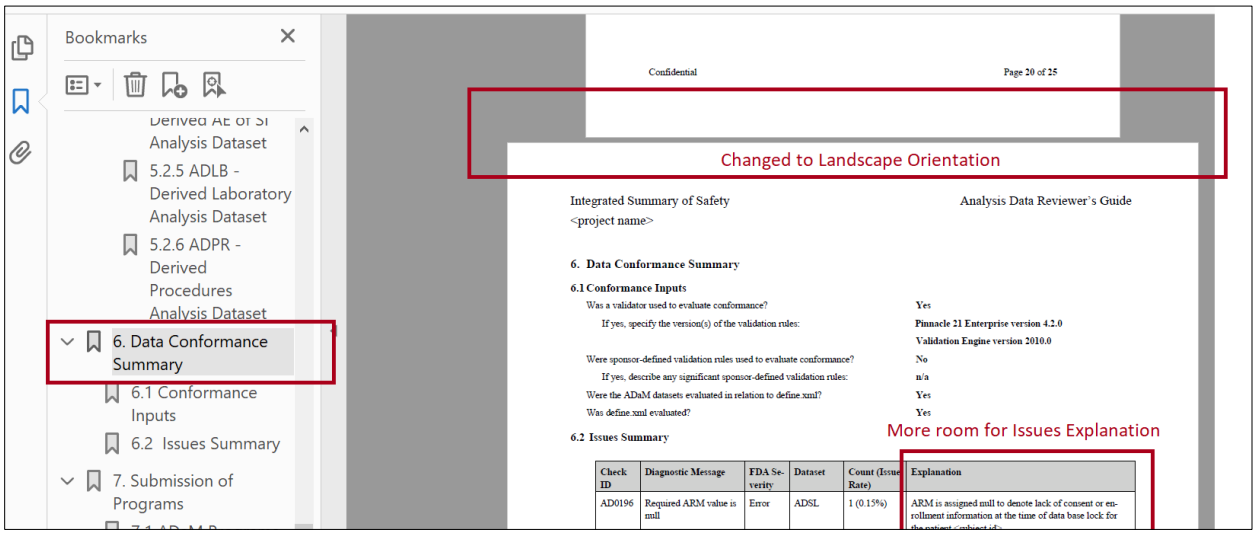

Figure 3: Changing Page Orientation for Data Conformance Sections

5. Clearly identify data sets that contain the primary efficacy endpoints in the "Analysis Data Sets" section of the ADRG by adding verbiage like "XX data set contains the primary efficacy endpoint variables (ORR, DOR, PSF, and OS)"
6. Consider using graphics, flowcharts and tables to document data dependencies and derivations. This may sometimes make complex information clearer compared to plain text.
7. In the cSDRG section "Appendix I: Inclusion/Exclusion Criteria", it will help to consolidate the rows with the same criteria. For example, write "A00 – A03" for protocol versions that have repeated criteria, instead of repeating all the information. This will help make the section concise, at the same time retaining all required information.

<project name> Clinical Study Data Reviewer's Guide

Protocol/ Amendment Version	Category	IETESTCD	Full Text of Criterion
A00 – A03	Inclusion	INCL003	Age 18 years or older.

Figure 4: Consolidating Inclusion/Exclusion Criteria in cSDRG

8. The issue explanation should be to the point with relevant details. Provide meaningful explanations in the P21 Issues Summary. Note that many P21 issues are more easily addressed when these checks are run throughout data set development well before final delivery. If P21 checks are not run until right before the submission itself, it is often much more challenging to go back and change study data sets without impacting reported results and locked-down analysis documentation.

   Examples:

   a. For P21 messages like "XX value not found in 'Frequency' extensible codelist", the explanations can be something like "The following sponsor values have been added to the extensible code list in define.xml: XX1, XX2, XX3."

   b. Any data-related issues should have been queried and reconciled during data cleaning efforts. For any open queries that generate P21 messages, the explanation should focus on the data cleaning and query resolution process, as shared by CDM. Just saying "Data not available" or "Data not cleaned" is less helpful.

   c. Issues like label inconsistencies between data sets, CDISC standards and define.xml, or variables not recommended for use, should be fixed in the data appropriately. For variables not recommended for use in SDTM, the variables should be included in SUPP-- instead of the parent domain. Even if some P21 issues don't look critical for analysis, it is a good practice to resolve as many issues as possible, so the reviewer's guide has only a handful of unresolved issues and which will inspire more confidence with agency reviewers in the quality of the data and standards.

   d. The explanation language should be consistent throughout the document. If explanations are being entered by multiple programmers, the language should be made consistent eventually. Consistent language in explanation will help reviewers to maintain a good flow of the review, rather than trying to understand different language and phrasing styles.

9. Consolidating and combining similar issues will make reviewer's guides short and concise.

   Examples: If issues have the same explanation, they can be consolidated.

PCORRESU value not found in 'Unit' extensible codelist
PCSTRESU value not found in 'Unit' extensible codelist
→ **PCORRESU/PCSTRESU** value not found in 'Unit' extensible codelist

SD0002	NULL value in ACTARMCD/ ACTARM/ARMCD/ARM variable marked as Required	DM	1 (0.91%)	Assigned missing ARM/ARMCD /ACTARM /ACTARMCD to denote lack of enrollment information at the time of data base lock.

10. Any extensive details like lookup tables should be added as separate supplemental documents linked to define.xml, instead of adding their content to reviewer's guides, which makes them

unnecessarily long and bulky. Reviewers can refer to the supplemental document through define.xml, if needed.

11. Specific variable derivations should not be in a reviewer's guide; those belong in define.xml. Some information like protocol details, list and description of data sets, list of core and critical variables, etc. can be repeated in the reviewer's guide for ease of review.
12. Reviewer's guides should be written such that there is no need for frequent back-and-forth navigation between the reviewer's guide, define.xml and protocol.
13. Run a spellcheck on the reviewer's guides to catch and correct obvious typos.

SUBMISSION PROGRAMS

FDA expects sponsors to share ADaM data set programs, more for reference purposes. It is also a good practice to submit TLF programs as agreed with FDA in initial submission package/structure meetings or communications. These programs can be limited to primary and secondary efficacy endpoints, and key safety TLF outputs. In many companies this could correspond to programs for topline or executive results produced within a few days after database lock.

Some tips on program submission:

1. Ideally, all pointers below are implemented before the final program run and output delivery. If any changes are necessary after the analysis is already closed out, adjusting a copy of the programs, rather than the source programs themselves, is recommended instead. In such cases, ensure the changes would not change the analysis results if the code were rerun.
2. Ensure there is no dead or redundant code in the submission programs.
3. The program headers can be updated to include only relevant information, like author name, purpose, input/source data, output, etc. Modification logs can be removed, as they are not necessary to be shared with FDA, and it will help make program header section cleaner.
4. The programs must be generously commented, to help understand various programming steps.
5. Any internal development comments, programmer names, reminders, etc. should be removed.
6. Consider submitting executable programs to FDA. This is optional unless explicitly agreed with FDA in any pre-submission communications.
    a. Submitting a separate executable program for primary endpoint(s) (with appropriate execution instructions) may address some of the FDA questions proactively and reduce potential Information Requests.
    b. If the list of TLF output programs to be submitted is small (e.g., 20 programs), consider making all programs executable. This can be done by writing selected TLF output testing programs free of external dependencies, macros, etc. These programs do not need to produce formal RTF outputs, rather can produce results in the SAS® LST result file, which may be sufficient for FDA review.

ADDITIONAL DEFINITIONS AND SUPPLEMENTAL DOCUMENTS

Additional definitions and supplemental documents are tremendously helpful for managing study specifications, and any special data management rules, that cannot be documented in detail in define.xml or reviewer's guides.

These documents can have extensive derivation rules and algorithms, any decision memos, sponsor lookup files, etc.

1. These documents can follow similar rules and formats as in reviewer's guides.

2. The documents must be written in structured format, with appropriate cover page, ToC, scope/purpose, appropriate sub-sections, numbering scheme (1, 1.x, 1.x.y…) and bookmarks.
3. As much as possible bullets should be avoided and numbering should be used. The numbering scheme helps with easy referencing.
4. If a table flows to the multiple pages, repeat the column headers on each page for ease of review.
5. Consider presenting some algorithms as graphics, flowcharts or tables. Sometimes it makes it easier to understand the complex algorithms and steps in the form of a graphics, flowchart, or table rather than writing in regular text.
6. Appropriate pagination should be done to avoid unnecessary white space, that may make the document seem bigger than it actually is.
7. Run a spellcheck on the documents to catch and correct typos.
8. In general, the document language should be simplified with appropriate details, that reviewers may find relevant and helpful.
9. As much as possible, files should be in PDF format (and follow eCTD guidelines and properties). When linked in define.xml, PDF files work best as they can be opened directly from the browser without needing to download or save (unlike Excel and other such files).
10. If multiple Excel files are planned to be submitted, like sponsor lookup files, they should be consolidated in one Excel file as individual, separate tabs and appropriately named. It will be more efficient for reviewers to review a single file.
11. Supplemental document file names should be shorter and relevant. File names with unnecessary detail like dates, prefix, suffix, company/team member names, that are more for internal use should be removed before submitting to FDA.

PRE-SUBMISSION AND INSPECTION READINESS

There are several activities that can be done proactively during the development phase, at the time of study analysis, data set and output programming and package preparation. This will not only have us more ready for submission but also for potential inspections, while at the same time ensuring fewer late-stage changes and giving us more time to think about what reviewer and inspectors would like to see.

PRE-SUBMISSION ACTIVITIES CHECKLIST

Sponsors can put together a checklist that can help ensure the submission package is complete with all required components.

TABLE 2: SAMPLE PRE-SUBMISSION CHECKLIST

Checklist before the Final Release (Pre-submission)	SDTM	ADaM	TLFs
Ensure specifications (including SDTM/ADaM/CT/VLM) are completely and accurately documented in define.xml, or additional definitions document	X	X	
All required submission components are included in the package: SDTM XPT, aCRF, cSDRG, SDTM define.xml, ADaM XPT, ADRG, ADaM define.xml, ADaM programs, TLF programs, supplemental documents, any other files.	X	X	

Checklist before the Final Release (Pre-submission)	SDTM	ADaM	TLFs
The Protocol, CT, CTCAE and Medical Dictionary (MedDRA, WHODrug) versions for pivotal study, ISS and ISE/SCE should match whenever applicable.	X	X	
There are no dead or redundant codes in the submission programs	X	X	X
The program headers are updated to include only the relevant information, modification logs are removed. The programs comments are appropriately documented.	X	X	X
Variable lengths are readjusted to maximum text length, to reduce size of the submission files.	X	X	
For the final rerun the correct sequence is followed for timestamp purpose and Inspection readiness. Example: SDTM-> Testing -> ADaM -> Testing -> TLF -> Testing.	X	X	X
The logs are clean. All error/warning/notice messages have been resolved unless they are expected messages.	X	X	X
The aCRF file has been flattened to avoid accidental edits/changes in annotations	X		
All PDF files included in submission package follow eCTD properties	X	X	
TS domain is consistent across all BLA/NDA studies and submission packages, especially for treatment information.	X		
EXTRT variable is consistent across all BLA/NDA studies and submission packages.	X		
Content is consistent across defines, RGs and datasets within the package. Example - protocol information, study name, etc.	X	X	X

INSPECTION READINESS

Inspection readiness is another aspect that sponsors can prepare proactively, and not just at the time of the inspection. Some pointers to be considered for inspection readiness:

1. If the study is randomized/blinded, the working folders should be set up to assign access to appropriate team members. Save the documents and evidence related to the access assignment process and ensure the restricted team is included in the appropriate user groups and access is tested for accuracy.
2. Programs and activities (production, testing: SDTM, ADaM, TLF) follow required guidelines and SOPs during the development phase.
3. Keep all documents up to date that are anticipated to be included in the submission, for example any memos, data cut specifications, lookup files, additional documentation documents, etc.
4. Submission activities follow applicable SOPs, training guides, Study Data Technical Conformance Guide, eCTD guidelines, etc.
5. Testing reports are generated and signed by respective team members.

6. Updated CVs are collected for all team members who worked on the filing.
7. All team member training records are up to date.

Knowing and being aware of these aspects upfront may help sponsors prepare for submission activities more mindfully and in a more informed manner.

CONCLUSION

Creating a quality submission package requires not just technical know-how, being aware of some small details that have been shared in this paper can go a long way to help make FDA reviewers' experience more pleasant. Throughout the lifecycle of these activities, keeping the reviewers' interests in mind can help make the submission package preparation process less challenging. It is important to follow all technical guidelines shared by FDA and various other industry tools and standards, to create an accurate, high-quality, and robust submission package. Some of these additional touch-ups can serve as icing on the cake in helping FDA reviewers.

REFERENCES

Study Data Technical Conformance Guide: https://www.fda.gov/media/143550/download

eCTD Guidelines: https://www.fda.gov/media/135373/download

CDISC Standards and Guidelines: CDISC.org

SDTM aCRF Guideline: https://wiki.cdisc.org/download/attachments/113589261/aCRF_Guideline_v1-0_20201120_publish.pdf?version=1&modificationDate=1606251671510&api=v2

PHUSE Reviewer's Guide Templates: https://advance.phuse.global/display/WEL/Deliverables

ACKNOWLEDGMENTS

I would like to thank my leaders and team for their constant support.

CONTACT INFORMATION

Your comments and questions are valued and encouraged.

Contact the authors at:

Shefalica Chand
Seagen Inc.
21823 - 30th Drive S.E.
Bothell, WA 98021
425-527-4866
schand@seagen.com