

An eSubmission Tracking Tool for Individual Components and Timelines

Todd Case

Vertex Pharmaceuticals, Boston, USA

ABSTRACT

There are at least as many ways to create an eSubmission as there are companies that do it. Given all the different components required for an eSubmission, this paper will demonstrate a way to manage, organize and create a timeline for all the deliverables, including all of the specific eSubmission components.

INTRODUCTION

Within a single (often large) company, sometimes even different therapeutic areas (TA's) use a different method (e.g., legacy programs with sNDA/BLAs may use an older system or vendor that they used with the original NDA/BLA while newer TA's may use a new custom built solution). Given all the different eSubmission components (e.g., annotated and bookmarked Case Report Form (aCRF), links to aCRF page numbering from Define-xml, executable table programs, ADaM programs, any external files, etc.) that are required for an eSubmission this paper will focus on a tool to manage, organize and create a timeline for the granular and overall information.

The benefit of this tool is that anyone, from the individual contributor who is responsible for creating/compiling NCI controlled terminology or executable table programs for a study, to the VP who is looking at the tracker every week, can get a transparent and clear idea of the percentage of the submission that is complete, where problem areas may lay, where things may routinely go faster than expected, etc.

NOT AN AFTERTHOUGHT

The days of eSubmission packages being started after the final Database Lock (DBL) to be included in the eSubmission are long over. eSubmission packages are something that should, and can, be prepared as early as you start writing your SDTM and ADaM specifications. If the specifications are written in a submission frame of mind, and your eSubmission tool is based on your specification, the amount of work done up front should be in direct proportion with the amount of study work completed, as depicted in Figure 1, below:

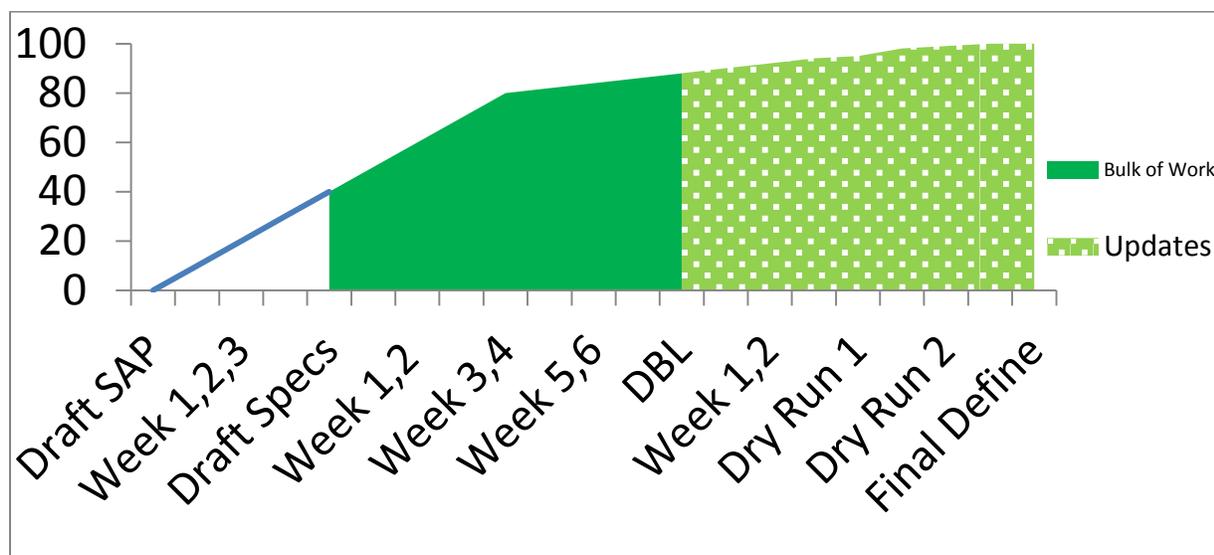


Figure 1: Example of Percentage of eSubmission Work Complete by Major Study/Submission Milestones¹

PRE-REQUISITES

In order to understand and get the most out of this paper, it's imperative to re-iterate the goal - to clearly describe what all of the elements of an eSubmission are and how these can be overlaid with a timeline to fit any submission project. From this perspective, it's Process first, then employing technology to implement the process.

But, before we can discuss the process, we should take a step back and understand each component within an eSubmission package and where it's "home" is within the electronic Common Technical Document (eCTD).

The folder structure for the data (e.g., Biometrics Deliverables) within the eCTD is depicted in Figure 2, as follows:

Figure 1: Folder Structure for Study Datasets

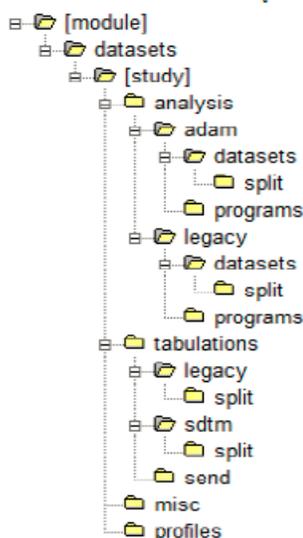


Figure 2: Directory Structure of an eSubmission Package: Module 5 (Clinical Study Reports) of the eCTD
<https://www.fda.gov/downloads/forindustry/datastandards/studydatastandards/ucm384744.pdf>

The Pre-requisites and individual components are as follows:

- ▶ Annotated CRF
 - Bookmarked (by Visit and Domain)
- ▶ Data Standards Catalog/Study Data Standardization Plan**
- ▶ Datasets (XPT including PC/PP, compressed, external data e.g., LB Conversion Factors, etc.)
- ▶ Metadata Creation*
- ▶ National Cancer Institute (NCI) Controlled Terminology
- ▶ SDRG
 - OpenCDISC Validation Report and Review
- ▶ ADRG
 - OpenCDISC Validation Report and Review
- ▶ ADaM Programs
- ▶ Executable Table Programs
- ▶ Datasets (XPT, compressed, external data, etc.)
- ▶ Define-XML (by study for both SDTM and ADaM)

* For input into the tool that creates Define-xml

**See reference: http://www.phusewiki.org/wiki/index.php?title=Study_Data_Standardization_Plan_%28SDSP%29

Each of these components can be placed in one of the following sub(directories), as shown in Figure 3:

Folder Name	Folder Level	Description/Contents
 [module]	1	Refers to the eCTD module in which study data is being submitted. Name this folder m4 for nonclinical data and m5 for clinical data. Do not place files at this level.
 datasets	2	Resides within the module folder as the top-level folder for study data (nonclinical or clinical) being submitted for the specified module (m4 or m5). Do not place files at this level.
 [study]	3	Name this folder with the study identifier or analysis type performed (e.g., study123, iss, ise). Do not place files at this level.
 analysis	4	Contains folders for analysis datasets and software programs; arrange in designated level 6 subfolders. Do not place files at this level.
 adam	5	Contains subfolders for ADaM datasets and corresponding software programs. Do not place files at this level.
 datasets	6	Place ADaM datasets in this subfolder.
 split	7	Place any split ADaM datasets in this subfolder.
 programs	6	Place software programs for ADaM datasets, tables and figures in this subfolder.
 legacy	5	Contains legacy formatted analysis datasets and corresponding software programs. Do not place files at this level.
 datasets	6	Place legacy analysis datasets in this subfolder.
 split	7	Place split legacy analysis datasets in this subfolder.
 programs	6	Place software programs for legacy analysis datasets, tables and figures in this subfolder.
 misc	4	Place miscellaneous datasets that don't qualify as analysis, profile, or tabulation datasets in this subfolder. This subfolder was formerly named "listings".
 profiles	4	Place patient profiles in this subfolder.
 tabulations	4	Contains subfolders for tabulation datasets. Do not place files at this level.
 legacy	5	Place legacy (non-standardized) tabulation datasets in this folder.
 split	6	Place any split legacy tabulations datasets in this subfolder.
sdtm	5	Place SDTM tabulation datasets in this subfolder. Should only be used in m5 for clinical data.
split	6	Place any split SDTM files in this subfolder.
send	5	Place SEND tabulation datasets in this subfolder. Should only be used in m4 for animal data.

Figure 3: Sub-Directory Structure of an eSubmission Package

<https://www.fda.gov/downloads/forindustry/datastandards/studydatastandards/ucm384744.pdf>

PROCESS

Now that we've talked about the what, it's time to discuss the process - laid out as follows, in the following order and with the necessary pre-requisites:

- 1) Having a team dedicated to working in a 'submission state of mind' – where specifications are created in such a way as to think about the future reviewers when they are written
 - a) This is absolutely critical – if the study team is in 'too much of a hurry' to enter submission comments early on it's going to take a significant amount more time to do this in the period right before the eSubmission, when teams are finishing last-minute analysis, documents (including define-xml), etc. are being reviewed.
- 2) Specifications that feed into an eSubmission/Define-xml Tool
 - a) Again, very critical – if you are writing specs but need to retrofit or modify these to fit into a canned eSubmission software or package it's very likely it's not going to be the seamless process that having your dataset specifications feeding directly into your eSubmission tool will be.
- 3) An eSubmission/Define-xml Tool that is robust enough to be able to incorporate data that is compliant with the spec as well as any form of legacy data
 - a) Although as mentioned above it's not ideal to have to retrofit specifications into your define-xml tool, actually having the ability to do this is necessary if teams are going to be able to create their define-xml(s)!
4. Define who WHO is responsible for Deliverable and WHEN each Deliverable is due
5. Define who WHO is responsible for Deliverable QC and WHEN QC of each Deliverable is due
6. Number 4 and 5, while seeming to be commonsense, are indispensable. With so many pieces involved it's absolutely imperative that each of these [components] have a person assigned for development and QC and a due date for both. These dates can be color coded for review by upper management, as demonstrated in Figure 4, below:

aCRF – Jane S	Bookmarked aCRF – John S	QC – Jane D	ADaM Programs – John D	QC – Tim J	XPT Datasets (Compressed)- John H	QC- Ani T	Executable Table Programs - Rory A	QC- JR	...
23Dec2016	01Jan2017	05Jan2017	02Jan2017	06Jan2017	10Jan2017	15Jan2016	18Jan2016	25Jan2016	

Figure 4: A Dashboard of eSubmission Components – 19Jan2016 (example)

TIMELINES/MILESTONES

Timelines and milestones can differ widely based on the scope (rare disease vs a cardiovascular therapy), and nature (BLA vs. sBLA) of the submission. In addition, the size of the company and whether there is a pre-existing process in place for creating define-xml or whether it is outsourced can also be major factors in terms of how realistic timelines are created. Given a simple example of a project that has a limited number of studies and there is pre-existing tool for creating define-xml based on the SDTM and ADaM specifications, let's assume that the submission will be out of the door six weeks after DBL, with the major milestones as follows: (1) draft Statistical Analysis Plan (SAP), which facilitates the creation of SDTM/ADaM specifications, (2) the weeks between draft SAP and Database Lock (DBL), when the majority of the define-xml work gets done (by finalization of the SDTM/ADaM specifications – again refer to Figure 1, (3) Database Lock, after which all the Tables, Listings and Figures get created, QC'd reviewed and finalized, and (4) two Dry Runs for review of the define-xml's (let's say the 1st Dry Run is three weeks after DBL, the second is 4.5 weeks after DBL, and the Final Define-xml is uploaded six weeks after DBL. It's a

simplified version of a submission, but one can see how it can easily be expanded to larger submissions (e.g., each study can be a row in the Tracker).

CONCLUSION

The purpose of this paper is to demonstrate to the reader a way to understand (and track) every single component of a Biometrics eSubmission package, from the aCRF all the way to define.xml. Once each component has been identified, a programmer needs to be identified to generate the component, QC the component (if applicable) and a date needs to be determined for both development and QC. These then need to be incorporated into an overall timeline; with the project lead will use to confidently track, monitor and report on the status of an eSubmission.

ACKNOWLEDGMENTS

I'd like to acknowledge all my colleagues and management at Vertex Pharmaceuticals for their dedication to standards and quality.

REFERENCES

- <https://www.cdisc.org/education/course/introduction-define-xml>.
- http://www.phusewiki.org/wiki/index.php?title=Study_Data_Standardization_Plan_%28SDSP%29
- <https://www.fda.gov/downloads/forindustry/datastandards/studydatastandards/ucm384744.pdf>
- Fitzmartin, Ron and Hussong, Ginny. "Required Electronic Submissions to CBER/CDER". *October 8, 2015*. Available at <http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/SmallBusinessAssistance/UCM467501.pdf> .
- "Providing Regulatory Submissions In Electronic Format — Standardized Study Data Guidance for Industry". US Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER). December, 2014. Available at <http://www.fda.gov/downloads/Drugs/.../Guidances/UCM292334.pdf> .

CONTACT INFORMATION

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

Name: Todd Case
Enterprise: Vertex Pharmaceuticals
Address: 50 Northern Avenue
City, State ZIP: Boston, MA 02210
Work Phone: 617 961 7907
E-mail: todd_case@vrtx.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.