Preparing FDA Submission Data Packages

Yuguang Zhao, MS
Senior Vice President, Programming and Development

Tiepu Liu, MD, PhD
President of Global Biometrics

ClinChoice

Contents

Executive Summary

The U.S. Food and Drug Administration (FDA) is now requiring that studies initiated after December 17, 2016 be submitted in New Drug Applications (NDAs), Biologics License Applications (BLAs) and Abbreviated New Drug Applications (ANDAs) according to accepted Clinical Data Interchange Standards Consortium (CDISC) standard electronic formats including the Standard for Exchange of Nonclinical Data (SEND), the Study Data Tabulation Model (SDTM), and the Analysis Data Model (ADaM). Studies initiated after December 17, 2017 will be required in the same formats for Investigational New Drug Applications (INDs).[1]

The FDA is responsible for protecting the public health by ensuring the safety, efficacy, and security of human and veterinary drugs, biological products, and medical devices; and by ensuring the safety of our nation's food supply, cosmetics, and products that emit radiation. FDA is responsible for advancing the public health by helping speed innovations that make medical products more effective, safer, and more affordable and by helping the public get the accurate, science-based information they need to use medical products and foods to maintain and improve their health.[2]

The implementation of electronic data standards advances the agency's efforts by enabling more timely and effective regulatory reviews of submissions and their corresponding data. The FDA has clearly stated, "Data that are not standardized diminish the Agency's ability to review the data efficiently, resulting in manual, labor-intensive processes and inherent inefficiencies in the review. They also limit the ability to automate some routine analyses."[3] The standards open up a world of opportunity in terms of exchanging, reviewing, and sharing data. However, for some sponsors, they prompt numerous questions about how to plan for and submit electronic study data packages to the FDA. This paper will describe the evolution of relevant standards and provide recommendations for best practices in planning and preparing FDA submission data packages.

History

In order to understand current study data standards, it is helpful to know a little about the history of those standards, what has impacted their evolution, and the reasons why they have been implemented.

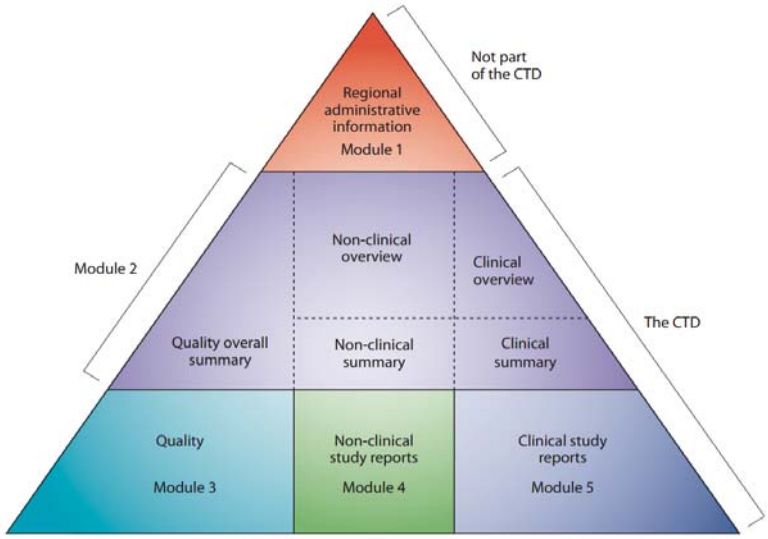
On March 20, 1997, the FDA published the Electronic Records and Electronic Signatures regulation (FDA 21 CFR Part 11). This regulation was significant because it signaled a real, practical shift toward electronic submissions and electronic standards and enabled applicants to voluntarily submit parts of their regulatory submissions in electronic format without requiring a backup paper copy. The regulation also spurred additional guidances and initiatives to support electronic submissions.

In April of 1998, CDER issued a draft guidance for the industry, Providing Regulatory Submissions in Electronic Format — NDAs[4]. This draft guidance built on the September 1997 guidance with information on submitting a complete archival copy of a New Drug Application (NDA) in electronic format. Later in 1998, the FDA Center for Biologics Evaluation and Research (CBER) published a draft guidance that focused on assisting applicants in submitting electronic content in the Biologic License Application (BLA) or Product License Application (PLA) and Establishment License Application (ELA) submission processes[5].

In 1999, the FDA CDER and CBER published Guidance for Industry Providing Regulatory Submissions in Electronic Format — General Considerations[6]. This **joint** guidance captured the common standards and considerations for all electronic submissions and was followed by guidances that addressed specific submission types (NDAs, BLAs, etc.).

As the FDA began to accept regulatory submissions in electronic format, a broader harmonization initiative brought the common technical document (CTD) to the international stage. The CTD was developed by a working group within the International Conference on Harmonization (ICH)[7]. It defined five discipline-specific modules for a regulatory submission. The first module is region specific while modules 2 through 5 are common for all regions.

Modular Structure of Common Technical Document[8]



The CTD triangle. The Common Technical Document is organized into five modules. Module 1is region specific and modules 2, 3, 4 and 5 are intended to be common for all regions.

This initiative sought to decrease the time, effort, and redundancies of submitting regulatory content to different health authorities around the world by identifying a common structure and set of content that would be accepted by all ICH members, including the US, Japan, and the European Union. As the CTD matured, the electronic Common Technical Document (eCTD) emerged to support electronic submissions.

As electronic submissions continued to make progress, the Clinical Data Interchange Standards Consortium (CDISC) emerged, addressing the data component of regulatory submission standards. CDISC is a global, open, multidisciplinary, non-profit organization that has established standards to support the acquisition, exchange, submission, and archive of clinical research data and metadata. The CDISC mission is to develop and support global, platform-independent data standards that enable information system interoperability to improve medical research and related areas of healthcare.[9]

In July 2004, the FDA first referenced the CDISC Study Data Specification in an eCTD guidance. Then, in December of 2006, the agency announced its intention to make the CDISC Study Data Tabulation Model (SDTM) required by regulation.

In May 2015, the FDA issued the guidance titled, Providing Regulatory Submissions in Electronic Format — Certain Human Pharmaceutical Product Applications and Related Submissions Using the eCTD Specifications (revision 3)[10].

All of these international initiatives grew from a need for electronic standards. Applicants were initially using no standards or developing their own. Even within the same company, standards would vary and this scenario required FDA medical and statistical reviewers to learn each company-specific or study-specific approach in order to review an NDA or BLA submission. This presented a significant challenge. Ramp-up time was required for each project review and for each new reviewer added to a project. The added time and resource costs started to pile up and had a negative impact on review processes and consequently on the time required to get new therapies to patients.

The continued collaboration of ICH, the FDA, and CDISC has enabled the development of electronic submission standards supported by technical documentation and specifications that are constantly being refined, updated, and implemented.

CDISC and FDA: Working Together

The FDA and CDISC have a long history of collaboration. Mutual goals have enabled the organizations to promote, implement, and refine electronic data standards. CDISC teams typically consist of volunteers from across the industry focused on the following:

- SDS – Submission Data Standards
- ADaM – Analysis Data Model
- SEND – Standard for Exchange of Non-clinical Data
- CDASH – Clinical Data Acquisition Standards Harmonization

The teams include FDA observers and are each focused on progressing specific standards models, including the development of supporting documentation such as implementation guides.

The Food and Drug Administration Safety and Innovation Act was signed into law in July 2012. This law expanded the agency's authority to collect user fees to fund product reviews and promote innovation with the intent of speeding patient access to safe and effective products. It also includes the fifth reauthorization of the Prescription Drug User Fee Act (PDUFA V). The production of guidances for the industry on the use of CDISC data standards for the electronic submission of study data is included in FDA's PDUFA goals.[11]

In December 2014, the FDA issued Providing Regulatory Submissions in Electronic Format — Standardized Study Data[12] along with four technical specifications. This binding guidance describes the requirements for an electronic submission of standardized clinical and nonclinical study data that use the standards specified in the Data Standards Catalog (Catalog). Technical specifications are referenced in this guidance document but are provided as standalone documents so they can be updated periodically. They include the list below and can be accessed on the FDA website[13]:

- Data Standards Catalog
- Study Data Technical Conformance Guide
- FDA Specific SEND Validation Rules
- FDA Specific SDTM Validation Rules

The Data Standards Catalog lists all supported and required versions of standards, their uses, and start and end dates for FDA support of each standard. It addresses exchange formats, study data, and controlled terminology standards that the FDA can process, review, and archive.

The Study Data Technical Conformance Guide replaced the Study Data Specifications document and CDER's Common Data Standards Issues document. It supplements the guidance and provides specifications and technical recommendations on submitting standardized data using those standards listed in the FDA Data Standards Catalog.

The binding guidance also notes:

- The Agency may **Refuse-To-File** (RTF) for NDAs and BLAs, or **Refuse-To-Receive** (RTR) for ANDAs if an electronic submission does not submit study data in conformance with the required standards specified in the Catalog.
- After the publication of this guidance, all studies with a start date **24 months** after the publication date must use the appropriate FDA-supported standards, formats, and terminologies specified in the Catalog for NDA, ANDA, and certain BLA submissions and 36 months for IND studies. (This means that most NDA/BLA and ANDA studies beginning after December 17, 2016 and December 17, 2017 are bound by this guidance, respectively.)
- If the application is not submitted electronically, the sponsor may receive an RTF or RTR.

The FDA's electronic submission guidances and regulations, the ICH eCTD specification, and the CDISC data standards form the foundation for sponsors to submit regulatory submissions and supporting data in electronic format that sustains the receipt and review of information by the FDA.

Standardized Study Data Package

A standardized study data package will include a SEND data package, SDTM data package, and an ADaM data package. Following is a brief list of the contents of each package:

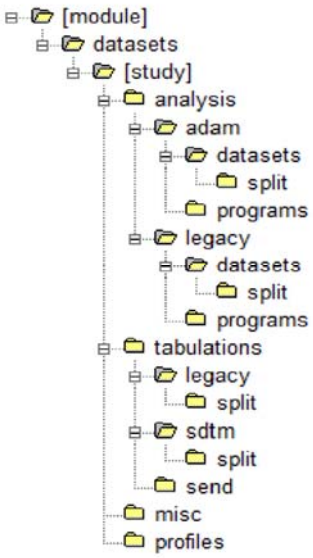
SEND Data Package	SDTM Data Package	ADaM Data Package
• acrf.pdf • .xpt files • Define.xml	• acrf.pdf • .xpt files • Define.xml	• .xpt files • Define.xml

• Study Data Reviewer's Guide (nSDRG.pdf)	• Study Data Reviewer's Guide (cSDRG.pdf)	• Analysis Data Reviewer's Guide (ADRG.pdf)

The Study Data Technical Conformance Guide [14] illustrates the standard folder structure for study datasets along with naming conventions, and a description of the contents as seen in the following graphic:

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 1: Folder Structure for Study Datasets



Planning with the End in Mind

The complexity involved in the design and creation of submission data packages can be a daunting task. There are several ways to simplify this endeavor with proper planning and early collaboration.

Study Data Standardization Plan (SDSP)

The December 2014 FDA Guidance identified the need for a Study Data Standardization Plan (SDSP). The plan describes the data standardization approach for studies within a development program. The development of an SDSP can help to facilitate internal discussions on the standards approach, provides a means of tracking discussions and agreements with the FDA documents on the rationale for legacy data conversion strategies as well as ongoing studies that support data pooling.

In developing the SDSP, sponsors should identify what studies will be submitted in the required standards. Studies that will be submitted in legacy study data formats should also be identified. It is important for sponsor teams to discuss the rationale and approach for each study early in the process to ensure that the data will be presented in the submission to support the regulatory review in the best way possible. If it is determined that a conversion is required, it is better to know early in the process and to plan and execute the conversion project with adequate time for quality checking and internal review.

The Pharmaceutical Users Software Exchange (PhUSE) is an independent, not-for-profit organization run by volunteers that has developed an SDSP template and implementation guide[15]. Both documents have been sent to the FDA for review. These documents provide a very general framework, or foundation, for sponsors to begin to compile their study data standardization information.

Conversion Decisions

Sponsors should discuss legacy data conversions as early in the process as possible. However, traceability is critical since the Clinical Study Report (CSR) may be based on legacy data. Key data verification with SDTM and ADaM datasets will also be necessary. This will also impact the

sponsor's data pooling strategy for the Integrated Summaries of Safety and Efficacy (ISS/ISE). Determine and document if you will pool at SDTM or ADaM level, and identify the SDTM and ADaM versions, WHO drug dictionary, and MedDRA dictionary versions.

Agency Communications

Standard sponsor/FDA meetings create opportunities to discuss the data standardization plan and the overall approach with the FDA prior to submission. The plan should be an official agenda topic for a scheduled agency meeting. Sponsors may start discussions at the pre-IND stage. For INDs, the plan should be located in the general investigational plan. CDER and CBER have jointly published recommendations for preparing the plan on the FDA website[16]. For Sponsors working toward an NDA, the plan should be submitted no later than the End-of-Phase-II Meeting. The pre-NDA/BLA meeting is considered too late in the process to allow strategy changes.

eCTD Test Submission

As a sponsor prepares to submit an electronic regulatory submission to the FDA, they can submit a test submission in advance of the actual submission. This test submission allows the sponsor to prepare a sample set of data and documents for the FDA for a meaningful, comprehensive analysis that helps to ensure that the sponsor is able to submit according to specifications. The test submission is processed just like a real submission so sponsors can identify any technical conformance issues. The test submission content does not go through a regulatory review by the agency.

Benefits of Implementing Data Standards

Compliance is the most important benefit of implementing data standards. An RTF or RTR risks and delays years of research and development; and can be costly to remediate. Implementing the standards early and leveraging mechanisms such as the SDSP and the test submission help sponsors to develop and articulate their approach to data standardization for internal and external stakeholders.

By leveraging FDA supported standards, sponsors can also facilitate the review process. FDA reviewers are trained on current data standards so they will require less support in understanding the structure/format of CDISC-compliant data sets.

Finally, in 2014, CDISC published a business case[17] for data standards based on research conducted in collaboration with Gartner, Inc., a leading IT research and advisory company The research suggested that communication among project team members and partners was improved, a greater level of accuracy requiring less training was achieved, and decision making was simplified when data standards were applied. The implementation of data standards also facilitated the transfer of data between partners, while positioning the sponsor to select from a wider choice of tools and technologies.

Findings suggest that by using CDISC standards from the start, researchers can save 70-90% of time and resources during the Study Start Up stage (time to first patient enrolled), and ~ 75% of non-patient participation time during Study Conduct and Analysis. The business case also suggests that standardizing data can shave 2 years off of an average 12 year clinical development program.

Conclusion

A submission data package requires a significant amount of internal and external collaboration in order to identify appropriate data standards, to develop a solid rationale for the data standards approach, to gain internal consensus, and to implement any required conversions. The FDA is able to provide valuable input into this discussion and should be used as a resource in this planning stage. Open communication with the FDA will add value throughout the regulatory review process.

Reference

[1] Providing Regulatory Submissions in Electronic Format — Submissions Under Section 745A(a) of the Federal Food, Drug, and Cosmetic Act, Guidance for Industry, U.S. 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.
[2] US FDA Website: http://www.fda.gov/AboutFDA/WhatWeDo/, FDA Mission, accessed January 18, 2017.
[3] Providing Regulatory Submissions in Electronic Format — Standardized Study Data, Draft Guidance. U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), Center for Devices and Radiological Health (CDRH), February 2012.
[4] Guidance for Industry Providing Regulatory Submissions in Electronic Format — NDAs, Draft Guidance, U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER), April 1998
[5] Electronic Submissions of a Biologics License Application (BLA), Product License Application (PLA) /Establishment License Application (ELA) to the Center for Biologics Evaluation and Research (63 FR 29741, 6/1/98).
Electronic Submissions of Case Report Forms (CRFs), Case Report Tabulations(CRTs) and Data to the Center for Biologics Evaluation and Research (63 FR 29739; 6/1/98).
Pilot Program for Electronic Investigational New Drug (eIND) Applications for Biological Products (63 FR 29740; 6/1/98).
Instructions for Submitting Electronic Lot Release Protocols to the Center for Biologics Evaluation and Research (63 FR 29742; 6/98).
[6] Guidance for Industry Providing Regulatory Submissions in Electronic Format — General Considerations, U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER), Center for Biologic Evaluation and Research (CBER), January 1999.
[7] International Conference on Harmonization Website: http://www.ich.org/about/history.html, accessed, January 18, 2017
[8] International Conference on Harmonization Website: http://www.ich.org/products/ctd.html, accessed, January 18, 2017.
[9] Study Data Standards for Regulatory Submissions Position Statement, September 15, 2003, FDA Website: http://www.fda.gov/ForIndustry/DataStandards/StudyDataStandards/ucm368613.htm, Position Statement, accessed January 18, 2017.
[10] Providing Regulatory Submissions in Electronic Format — Certain Human Pharmaceutical Product Applications and Related Submissions Using the eCTD Specifications Guidance for Industry, U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), May 2015, Electronic Submissions, Revision 3.
[11] Study Data Standards for Regulatory Submissions Position Statement, September 15, 2003, FDA Website: http://www.fda.gov/ForIndustry/DataStandards/StudyDataStandards/ucm368613.htm, Position Statement, accessed January 18, 2017.
[12] Providing Regulatory Submissions In Electronic Format — Standardized Study Data, Guidance for Industry, U.S. 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, Electronic Submissions.
[13]Study Data Standards Resources, FDA Website: http://www.fda.gov/forindustry/datastandards/studydatastandards/default.htm, Important Notices, accessed January 18, 2017.
[14] Study Data Technical Conformance Guide, Technical Specifications Document, U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), November 2016.
[15] Study Data Standardization Plan, PhUSE Website: http://www.phusewiki.org/wiki/index.php?title=Study_Data_Standardization_Plan_(SDSP), Project Deliverables, accessed January 18, 2017.

[16] CDER/CBER Study Data Standardization Plan Recommendations, FDA Website: http://www.fda.gov/downloads/ForIndustry/DataStandards/StudyDataStandards/UCM447119.pdf, accessed January 18, 2017.

[17] 2014 CDISC Business Case, Executive Summary, CDISC Website: https://www.cdisc.org/system/files/all/article/PDF/2014%20Business%20Case_Executive%20Summary.pdf, accessed January 18, 2017.

About the Authors

Yuguang Zhao, MS
Senior Vice President of Programming and Development

Yuguang Zhao is the Senior Vice President of Programming and Development at ClinChoice. Mr. Zhao is responsible for providing strategic planning and corporate direction for continued growth of the company, providing scientific leadership and resource planning for statistical support on all projects, and ensuring the teams provide high quality statistical programming support for Phase I, II, III, and IV clinical trials. He has over 19 years of experience in the pharmaceutical industry and has been very active in industry working groups. He is used to be a member in the CDISC SDS team and has received FDA Leveraging/Collaboration Award for the participation of creation of SDTM and accompanying implementation guide in 2005. He has co-led the CDISC/FDA data integration pilot. For questions on preparing FDA Submission data packages, please contact Yuguang at: yuguang.zhao@clinchoice.com

Tiepu Liu, MD, PhD
President of Global Biometrics

Tiepu Liu, MD, PhD, is the President of Global Biometrics at ClinChoice. Dr. Liu has served as Director of Biostatistics at PPD, Executive Director of Statistics and Data Management at UBC, and a Senior Director of Biostatistics at Graceway Pharmaceuticals and The Medicines Company, before joining ClinChoice to lead the Biostatistics and Data Management functions. Dr. Liu has published 100 scientific papers and more than 100 abstracts and presentations. He has served on numerous scientific review committees and panels for the National Institute of Health (NIH) and other government agencies. Dr. Liu has led numerous clinical development projects, interacting with various regulatory agencies and Data Monitoring Committees. He has been active in CDISC standards and CFAST initiative. For questions on preparing FDA submission data packages, please contact Tiepu at: Tiepu.liu@clinchoice.com