accenture

Strategy | Digital | Technology | Operations

High performance. Delivered.

Accenture Accelerated R&D Services

Rethink Reshape Restructure…for better patient outcomes

Preparing ADaM and Related Files for Submission

Sandra Minjoe
Senior ADaM Consultant

Presentation Focus

- Material covered in the Dec 2014 FDA Binding Guidances
  - Providing Regulatory Submissions in Electronic Format — Submissions Under Section 745A(a) of the Federal Food, Drug, and Cosmetic Act
  - Providing Regulatory Submissions In Electronic Format — Standardized Study Data
- Other FDA documents referenced by these binding guidances
  - Data Standards Catalog
  - Study Data Technical Conformance Guide
- And an FDA CDER document
  - Technical Rejection Criteria for Study Data
- More specifically
  - Analysis data and related files as part NDA and most BLA submissions to FDA CDER and CBER

Presentation Content

- References are found at the bottom of the slide2
  - Internet locations for references are found at the end of the presentation
- Black or Green text is used for quoting or summarizing the referenced documents
- Purple text is used for my own recommendations

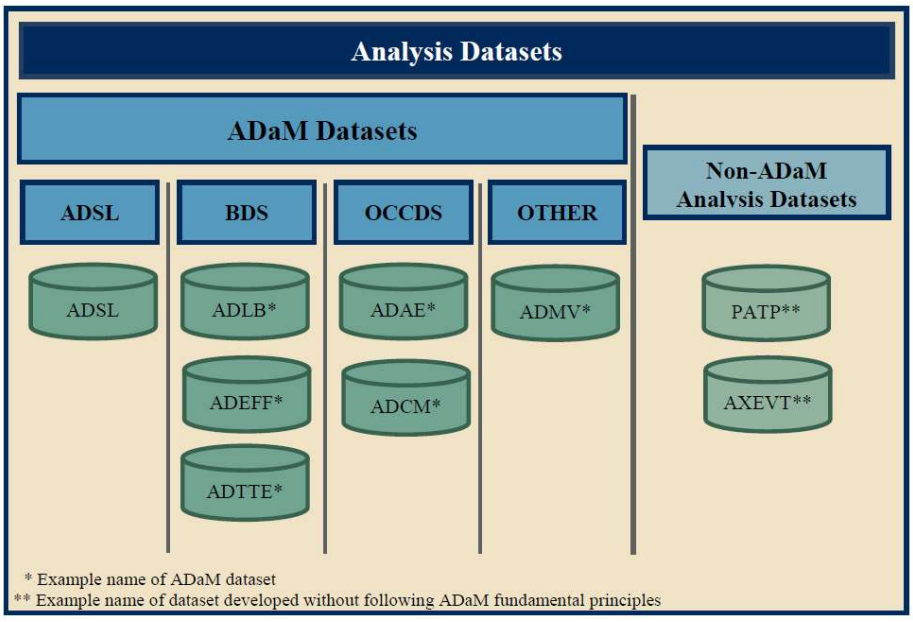
Preparing ADaM and Related Files for Submission

- What is Submitted?
  - Analysis and other related data
  - Analysis programs
  - Define files
  - ADRG

Analysis Datasets: Definitions*

- **Analysis dataset** – An analysis dataset is defined as a dataset used for analysis and reporting
- **ADaM dataset** – An ADaM dataset is a particular type of analysis dataset that either:
  1. Is compliant with one of the ADaM defined structures and follows the ADaM fundamental principles; or
  2. Follows the ADaM fundamental principles defined in the ADaM model document and adheres as closely as possible to the ADaMIG variable naming and other conventions
- **Non-ADaM analysis dataset** – A non-ADaM analysis dataset is an analysis dataset that is not an ADaM dataset. Examples of non-ADaM analysis datasets include:
  - An analysis dataset created according to a legacy company standard
  - An analysis dataset that does not follow the ADaM fundamental principles

 * Reference: ADaMIG v1.1, Feb 2016



 Reference: ADaMIG v1.1, Feb 2016

CDISC Dataset Standards for ADaM

- Dataset Standards from ADaM documents
  - ADSL
  - BDS
  - OCCDS if using ADaMIG v1.1
    - ADAE if using ADaMIG v1.0
- FDA also accepts standards from some CDISC TAUGs
  - Chronic Hepatitis C
  - Dyslipidemia
  - Diabetes
  - QT Studies
  - Tuberculosis

* References: ADaMIG v1.1, Feb, 2016
FDA Study Data Technical Conformance Guide, Jul 2016

Analysis Dataset Transport Files*

- The SAS Transport Format (XPORT) Version 5 is the file format for the submission of all electronic datasets
  - Why Such an "old" file format?
    - The SAS v5 XPORT is an open file format published by SAS Institute for the exchange of study data
    - Data can be translated to and from SAS v5 XPORT to other commonly used formats without the use of programs from SAS Institute (or any specific vendor)
- Submit one dataset per transport file
  - Transport file named the same as the dataset
  - Example: **adae** dataset -> **adae.xpt** SAS transport file
- Create SAS v5 XPORT files using SAS PROC COPY
  - Warning: SAS Transport files processed by the SAS CPORT cannot be reviewed, processed, or archived by FDA

* Reference: FDA Study Data Technical Conformance Guide, Jul 2016

Issues Using SAS v5 Transport Files

- When converting to SAS v5 transport file, newer features of SAS will be lost:
  - Longer variable names, labels, and text strings will be truncated
  - Newer formats will be lost (such as ISO8601 numeric date/time formats)
- Before submitting, ensure no data or formatting is lost
  - One possible process for testing:
    1. Create SAS datasets
    2. Create transport files from those original SAS datasets
    3. Convert transport files into new SAS datasets
    4. Compare the original SAS datasets with the converted SAS datasets to check for discrepancies

Other Dataset Requirements*

- The allotted length for each column containing text data should be set to the maximum length of the variable needed
  - In other words, don't artificially set to the 200 character length
- Datasets greater than 5 gigabytes (GB) in size should be split into smaller datasets no larger than 5 GB
  - Submit these smaller datasets, in addition to the larger non-split datasets, to better support regulatory reviewers
  - Split datasets are placed in a separate sub-directory labeled "split"
- But why even create analysis datasets larger than 5 GB?
  - Unlike SDTM, ADaM datasets can contain just what we need them to!
  - ADaM datasets can be "split" for ease of analysis, not just submission
  - Recommendation: create smaller datasets for analysis use
    - No splitting is needed for submission
    - Can also reduce analysis results program run time

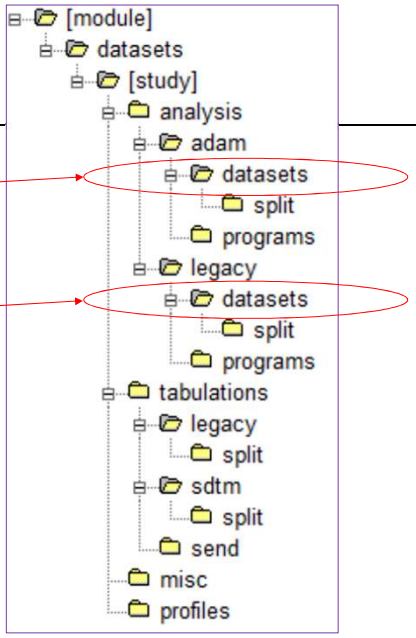
* Reference: FDA Study Data Technical Conformance Guide, Jul 2016

eCTD Data Folders*

Place ADaM datasets in this subfolder

Place legacy analysis datasets in this subfolder

Question: Should I split a study across these folders, with ADaM datasets in one folder and non-ADaM datasets in the other?



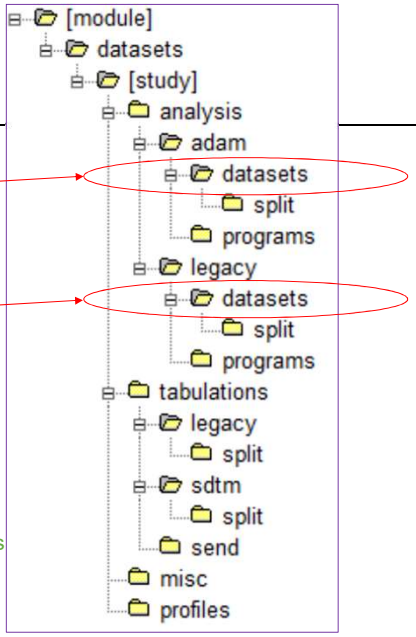
* Reference: FDA Study Data Technical Conformance Guide, Jul 2016

eCTD Data Folders *

Place ADaM datasets in this subfolder

OR

Place legacy analysis datasets in this subfolder



For ease of use with the define file and in the eCTD folder structure, all analysis datasets should be kept in one folder

- If a set of analysis datasets includes an ADaM-compliant ADSL dataset (as required for a CDISC-conformant submission), then the whole set of analysis datasets should be placed into the **adam** folder
- If not, the whole set of analysis datasets should be placed into the **legacy** folder.*

* Reference: ADaMIG v1.1, Feb 2016

ADaM or Legacy?

- ADaM will be required for studies starting after Dec 17, 2016
  - Recommended for studies NOW
- ADaM datasets should be used to create and support results in
  - Clinical study reports
  - Integrated Summaries of Safety (ISS)
  - Integrated Summaries of Efficacy (ISE)
  - Other analyses required for a thorough regulatory review
- ADaM v2.1/ADaMIG v1.0 are the only versions currently accepted
  - Check with Review Division if you want to use ADaMIG v1.1

- **References: FDA Study Data Technical Conformance Guide, Jul 2016**
  Providing Regulatory Submissions in Electronic Format — Standardized Study Data, Dec 2014
  FDA Data Standards Catalog v4.5.1, Sept 2016

Why Does FDA want ADaM?

- ADaM facilitates FDA review
- ADaM simplifies programming steps necessary for performing an analysis
- There are features built into the ADaM standard that promote traceability
  - From analysis results to ADaM datasets
  - From ADaM datasets to SDTM datasets
- FDA Reviewers are getting used to CDISC data
  - Tools
  - Training

*** Reference: FDA Study Data Technical Conformance Guide, Jul 2016**

Which Analysis Files Must I Submit?

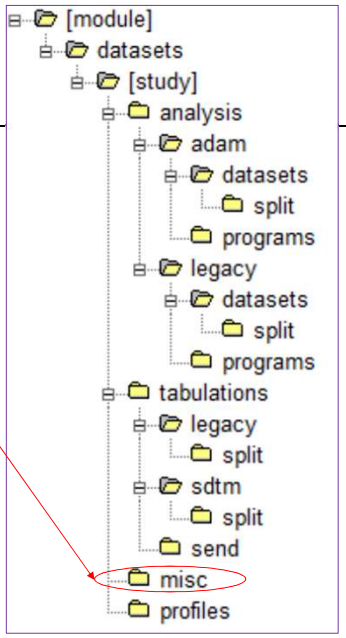
- From CDISC:
  - The sponsor determines the analysis datasets to be created
- From FDA:
  - Sponsors must submit **ADSL** for studies starting after 17DEC2016
  - Sponsors should submit ADaM datasets to support **key efficacy and safety** analyses
  - At least one dataset should be referenced as containing the primary efficacy variables
    - For example: ADEFF
- Sponsor can choose to not submit other datasets
  - This is a risk, so be prepared to submit later

*** References: FDA Technical Rejection Criteria for Study Data, Oct 2016**
FDA Study Data Technical Conformance Guide, Jul 2016
ADaM v2.1, Dec 2009

eCTD Data Folders *

Place miscellaneous datasets that don't qualify as analysis, profile, or tabulation datasets in this subfolder

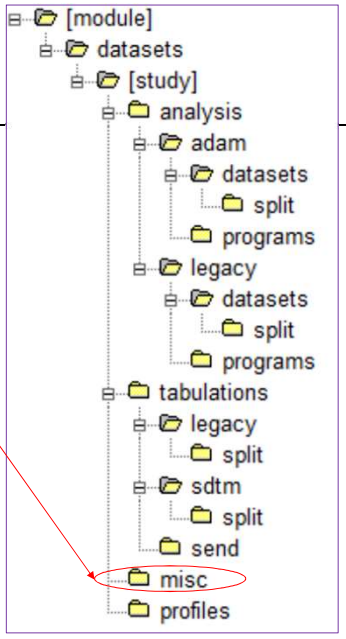
Question:
What kinds of datasets qualify as "misc"?



*** Reference: FDA Study Data Technical Conformance Guide, Jul 2016**

eCTD Data Folders *

Place miscellaneous datasets that don't qualify as analysis, profile, or tabulation datasets in this subfolder



Recommendation:

Include data not captured in SDTM but used to create ADaM datasets, such as:

- Look-up tables
- Deviations not collected via CRF

Preparing ADaM and Related Files for Submission

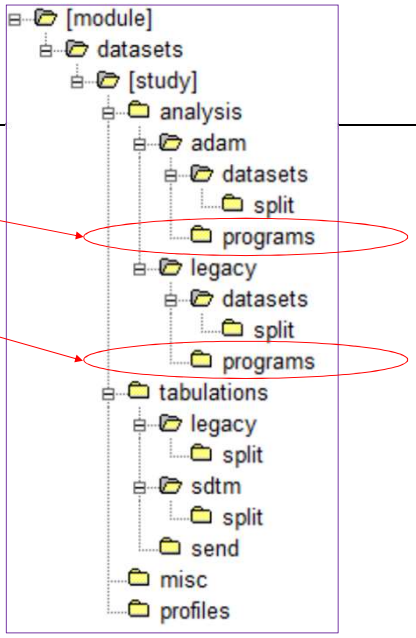
- What is Submitted?
  - Analysis and other related data
  - Analysis programs
  - Define files
  - ADRG

Analysis Programs

Place ADaM programs in this subfolder

OR

Place legacy programs in this subfolder



Provide programs used to create ADaM datasets, tables, and figures associated with primary and secondary efficacy

- Purpose is to understand the process and to confirm analysis algorithms
- Not necessary to submit the programs so that they can be directly run

* Reference: FDA Study Data Technical Conformance Guide, Jul 2016

Analysis Datasets and Programs

- Submit programs as ASCII text files (*.txt) or PDF (*.pdf)
  - FDA Reviewers may not use SAS
  - Example: **adtte.sas** submitted as **adtte.txt**

- Recommendations
  - Submit programs for at least:
    - Each dataset submitted
    - Key analyses
  - Be prepared to provide programs for every dataset and every analysis
  - Make the submitted programs as easy to read as possible
    - Include comments
    - Remove as many macros and macro variables as possible
    - May not need to include code that puts results on the table

* Reference: FDA Study Data Technical Conformance Guide, Jul 2016

Preparing ADaM and Related Files for Submission

- What is Submitted?
  - Analysis and other related data
  - Analysis programs
  - Define files
  - ADRG

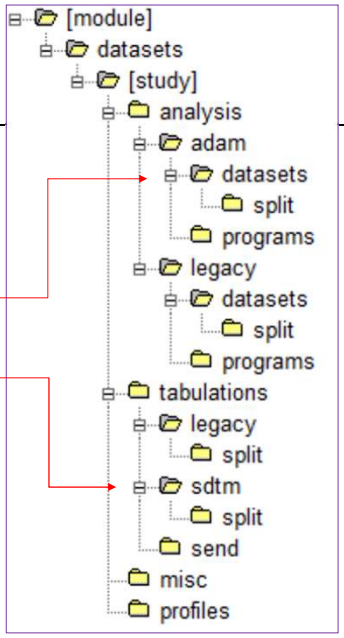
Define File

- The data definition (define) file
  - Describes the metadata of the submitted electronic datasets
  - Considered to be the most important part of the electronic dataset submission for regulatory review
- Submit the define file in XML format
  - Define.xml v2.0 is the preferred version
  - A printable define.pdf should be provided if define.xml cannot be printed
    - If define.xml v2.0 is used, there is no printing issue
- Can I submit define.xml v1.0 or even define.pdf?
  - Define.xml v1.0 is still in the FDA Data Standards Catalog, with support ending March, 2018
  - After this date, a waiver can be requested for an earlier standard

*** References: FDA Study Data Technical Conformance Guide, Jul 2016**
FDA Data Standards Catalog v4.5.1, Sept 2016
Providing Regulatory Submissions in Electronic Format — Standardized Study Data

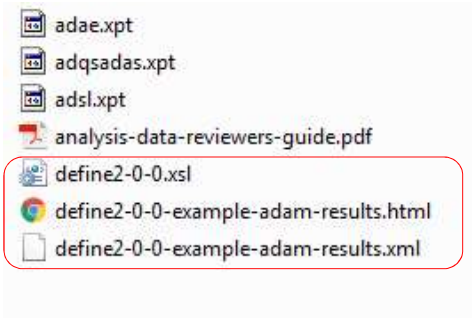
Define File Location

- Define file sits in the same folder as data
- Submit a separate define file for
  - ADaM
    - Or legacy analysis data
  - SDTM
    - Or legacy tabulation data
- Each data folder needs its own define



* Reference: FDA Study Data Technical Conformance Guide, Jul 2016

Example adam -> datasets Folder with Define



- Notice three “define” files
  - XML
  - HTML
  - XSL (Style sheet)

* Reference: Analysis Results Metadata Specification v.0 for Define-XML v2, Jan 2015

Define.XML Content

- Contents for ADaM define.xml can include
  - Dataset Metadata
  - Variable Metadata
  - Value Metadata, when appropriate
  - Results Metadata
    - Recommended for critical analyses
  - Controlled terminology and codes
  - Links to other documents
    - SAP
    - ADRG

- CDISC documents have examples of how to lay out a define

* Reference: Analysis Results Metadata Specification v.0 for Define-XML v2, Jan 2015
ADaM v2.1, Dec 2009

Preparing ADaM and Related Files for Submission

- What is Submitted?
  - Analysis and other related data
  - Analysis programs
  - Define files
  - ADRG

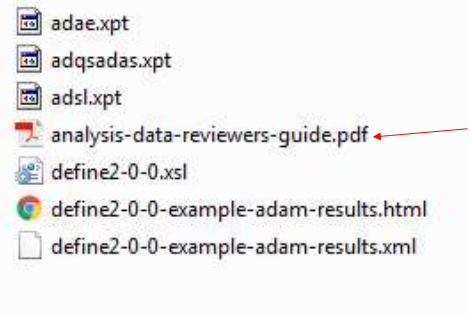
Analysis Data Reviewer's Guide (ADRG)

- An ADRG
  - Is recommended as an important part of a standards-compliant analysis data submission
  - Provides FDA reviewers with context for analysis datasets and terminology, in additional to what is presented within define.xml
    - Note: submission of an ADRG does not eliminate the requirement to submit a complete and informative define.xml file

- CSS (not CDISC) has developed a template ADRG
  - Examples can be found in
    - CSS ADRG zip file (which also includes the template)
    - Some CDISC define.xml v2.0 examples

* References: FDA Study Data Technical Conformance Guide, Jul 2016
Analysis Data Reviewer's Guide template and examples, Dec 2014
Analysis Results Metadata Specification v.0 for Define-XML v2, Jan 2015

ADRG Location



Belongs in the same folder as the datasets and define files

* References: Analysis Results Metadata Specification v.0 for Define-XML v2, Jan 2015
FDA Study Data Technical Conformance Guide, Jul 2016

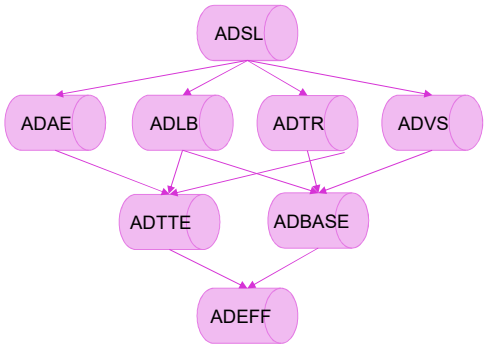
ADRG Standard Content

- Includes sections for
  - Acronyms, Standards, and Dictionaries
  - Data Source(s)
  - Protocol information
  - Analysis Variables of Interest
  - Dataset Processing
  - Data Conformance
  - Programs

* Reference: Analysis Data Reviewer's Guide template and examples, Dec 2014

ADRG Variables of Interest and Dataset Processing

- Variables of Interest
  - Core variables
  - Treatment variables
  - Imputation rules
  - Visit windowing



- Dataset Processing
  - Dataset dependencies
  - Intermediate datasets
  - Good place for a Flow Diagram to explain any complex data flows

* Reference: Analysis Data Reviewer's Guide template and examples, Dec 2014

ADRG Example: Conformance

6.1 Conformance Inputs

- Were the analysis datasets evaluated for conformance with CDISC ADaM Validation Checks?
  Yes.
- Were the ADaM datasets evaluated in relation to define.xml?
  Yes.
- Was define.xml evaluated?
  Yes.

The data and define were evaluated using the following OpenCDISC versions:

OpenCDISC version 1.4.1

CDISC Controlled Terminology Version 2011-07-22

6.2 Issues Summary

OpenCDISC Notices were evaluated for potential problems but are not listed here. The following is a summary of Error level messages. There were no Warning level messages.

Dataset(s)	Diagnostic Message and/or Check ID	Severity	Count/ Issue Rate	Explanation
ADAE	Neither AVAL nor AVALC are present in dataset	Error	1	The dataset is a hierarchical occurrence structure, the message is not relevant to this structure.

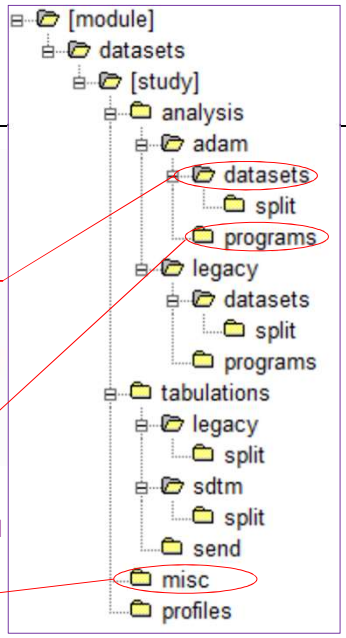
 * Reference: Analysis Data Reviewer's Guide example, Dec 2014

Recap

Datasets (SAS v5 transport)
- adae.xpt
- adqsadas.xpt
- adsl.xpt

ARDG
- analysis-data-reviewers-guide.pdf

Define files
- define2-0-0.xsl
- define2-0-0-example-adam-results.html
- define2-0-0-example-adam-results.xml



Programs for at least:
- Each dataset submitted
- Key analyses

Other data, such as:
- Look-up tables
- Deviations not collected via CRF

References used in building this presentation

- FDA guidance documents and catalog
  - http://www.fda.gov/ForIndustry/DataStandards/StudyDataStandards/default.htm
    - This is also where you'll find email addresses to ask questions to CDER/CBER
- FDA Technical Rejection Criteria
  - http://www.fda.gov/Drugs/DevelopmentApprovalProcess/FormsSubmissionRequirements/ElectronicSubmissions/ucm153574.htm
- CDISC ADaM and define.xml documents
  - http://www.cdisc.org/
- CSS ADRG documents
  - http://www.phusewiki.org/wiki/index.php?title=FDA_Working_Groups