

Preparing ADaM Datasets and Related Files for FDA Submission

Sandra Minjoe and Ragini Hari

PRA Health Sciences

MinjoeSandra@prahs.com and HariRagini@prahs.com



Presentation Focus

- Material covered in the Dec 2014 FDA Binding Guidance docs
 - **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 Guidance docs
 - **Data Standards Catalog**
 - **Study Data Technical Conformance Guide**
 - **Technical Rejection Criteria for Study Data**

More specifically

- Analysis data & related files in NDA & most BLA submissions to FDA CDER & CBER



Presentation Conventions

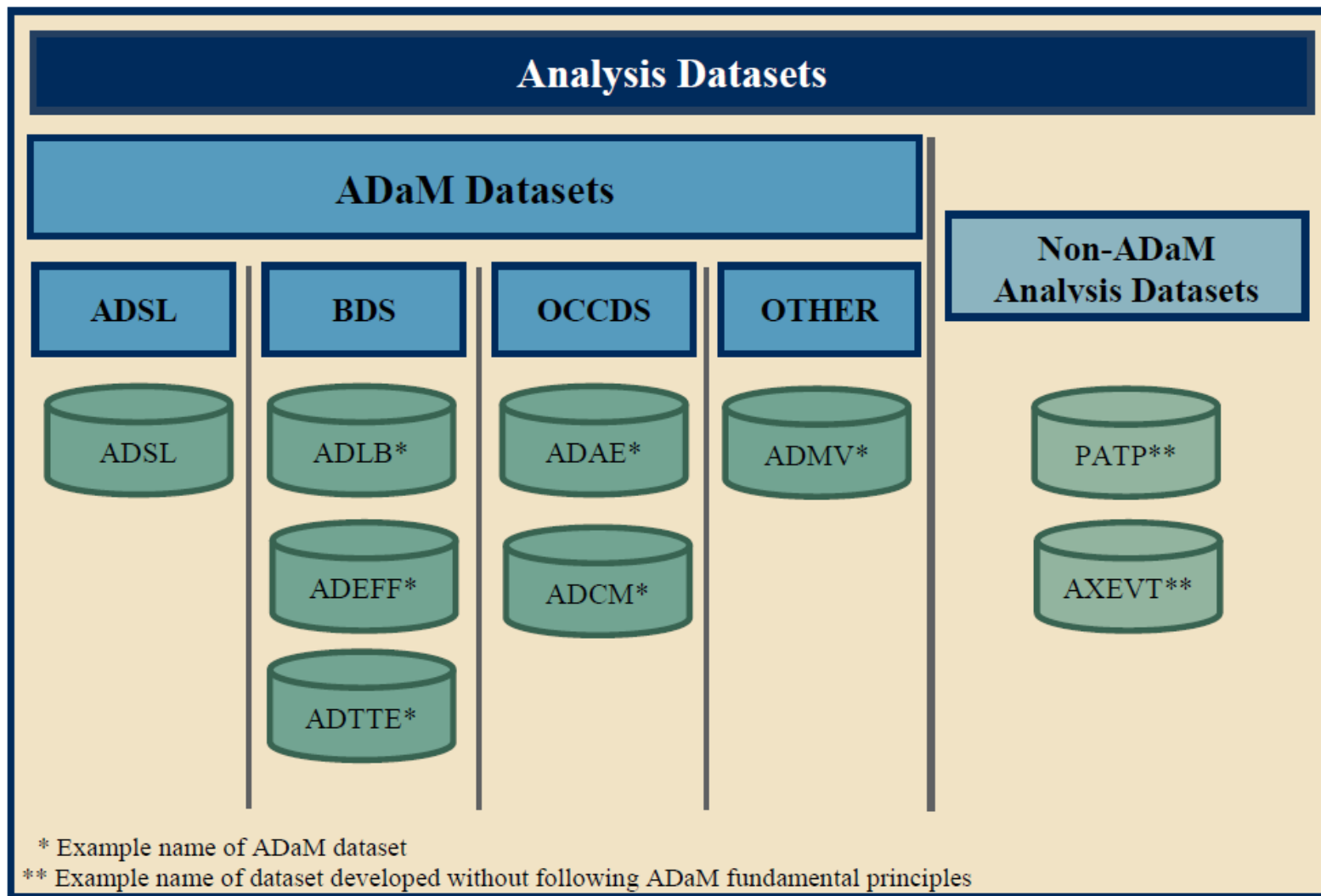
- References are found at the bottom of each slide
 - Internet locations for references:
 - <https://www.cdisc.org/standards/foundational/adam>
 - <https://www.fda.gov/ForIndustry/DataStandards/StudyDataStandards/default.htm#Catalog>
 - <https://www.fda.gov/Drugs/DevelopmentApprovalProcess/FormsSubmissionRequirements/ElectronicSubmissions/ucm153574.htm>
- **Black** or **Green** text is used for quoting or summarizing the referenced documents
- **Purple** text is used for my own recommendations

References for each slide are noted here

Preparing ADaM and Related Files for Submission

What is Submitted?

- ADaM and other related data
- Analysis programs
- Define files
- ADRG



Submitting ADaM & other Related Data



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 CDISC TAUGs
 - Chronic Hepatitis C, Dyslipidemia, Diabetes, Diabetic Kidney Disease, Ebola, Influenza, Kidney Transplant, Malaria, QT Studies, Rheumatoid Arthritis, Tuberculosis, Virology

References: ADaMIG v1.1, Feb 2016

FDA Study Data Technical Conformance Guide, sections 4.1.2 and 5.2, Mar 2018

Submitting ADaM & other Related Data



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, section 4.1.2.2, Mar 2018

Submitting ADaM & other Related Data



Format of Datasets Submitted

- 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 SAS
 - **Create SAS v5 XPORT files using SAS PROC COPY or SAS DATA STEP**
 - FDA can't use XPORT files created using SAS CPORT
 - Submit one dataset per transport file
 - Transport file named the same as the dataset (e.g. **adae.sas7bdat** -> **adae.xpt**)

Submitting ADaM & other Related Data



Issues Using SAS v5 Transport Files

- **When converting to SAS v5 XPORT, watch out for**
 - Truncation of long variable names, labels, and text strings
 - Lost formats (such as newer ISO8601 numeric date/time formats)
- **Before submitting, ensure no data or formatting is lost**



Submitting ADaM & other Related Data



Example to Check Transport Files

- **Sample code that can be used to test transport file readiness:**

```
data xptfile.adsl;  
  set adam.adsl;  
run;
```

Create transport file from original dataset

```
data new.adsl;  
  set xptfile.adsl;  
run;
```

Create new dataset from transport file

```
proc compare base=adam.adsl compare=new.adsl printall;  
  title "Comparison of adam.adsl (BASE) and new.adsl (COMPARE)";  
run;
```

Compare original dataset from new dataset



Submitting ADaM & other Related Data



Dataset Size Requirements

- Set each text variable length to the maximum needed
 - **In other words, don't artificially set to the 200 character length**
- Split datasets greater than 5 gigabytes (GB) in size into smaller datasets
 - Submit these smaller datasets, in addition to the larger non-split datasets
 - 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**

*References: FDA Study Data Technical Conformance Guide, sections 3.3.2 and 3.3.3, Mar 2018
ADaMIG v1.1, Feb 2016*

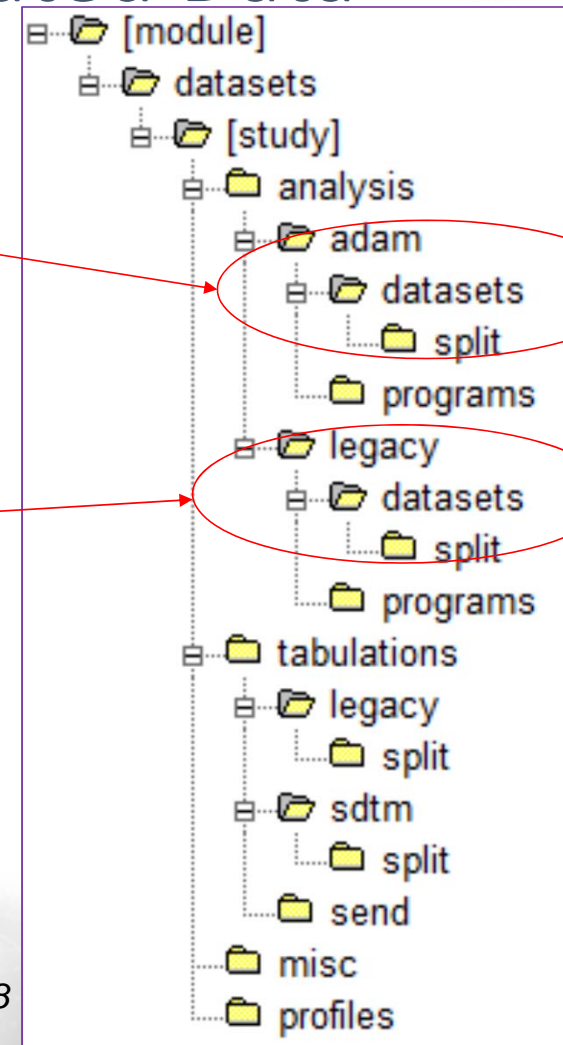
Submitting ADaM & other Related Data



Data Submission Location

If the study includes an ADaM-compliant ADSL, place the whole set of analysis datasets in this subfolder

Otherwise, place the study's whole set of analysis datasets in this subfolder



Submitting ADaM & other Related Data

Do I Have to Submit ADaM?

- ADaM is required for studies that start after Dec 17, 2016
 - **Recommended for studies NOW**

- Currently accepted versions of ADaM documents:
 - ADaM v2.1
 - ADaMIG v1.0
 - ADaMIG v1.1 (includes OCCDS v1.0, which is referenced in ADaMIG v1.1)

References: *FDA Study Data Technical Conformance Guide, Mar 2018*

Providing Regulatory Submissions in Electronic Format — Standardized Study Data, Dec 2014

FDA Data Standards Catalog v4.25.18, Apr 2018

Submitting ADaM & other Related Data

Which ADaM Datasets Must I Submit?

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

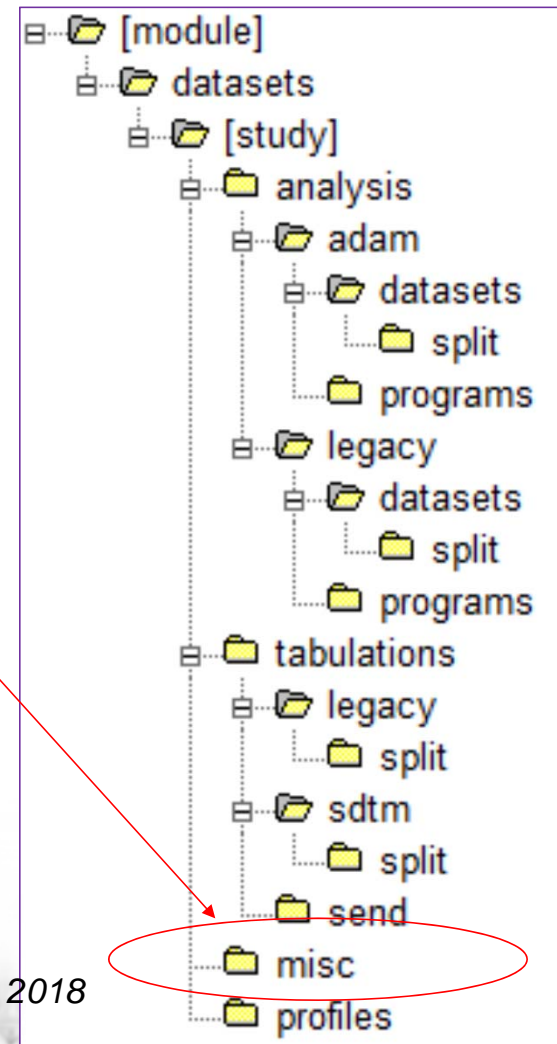
*References: FDA Technical Rejection Criteria for Study Data, May 2018
FDA Study Data Technical Conformance Guide, Jul 2016
ADaM v2.1, Dec 2009*

Submitting ADaM & other Related Data

Miscellaneous Data

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

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



Submitting ADaM & other Related Data



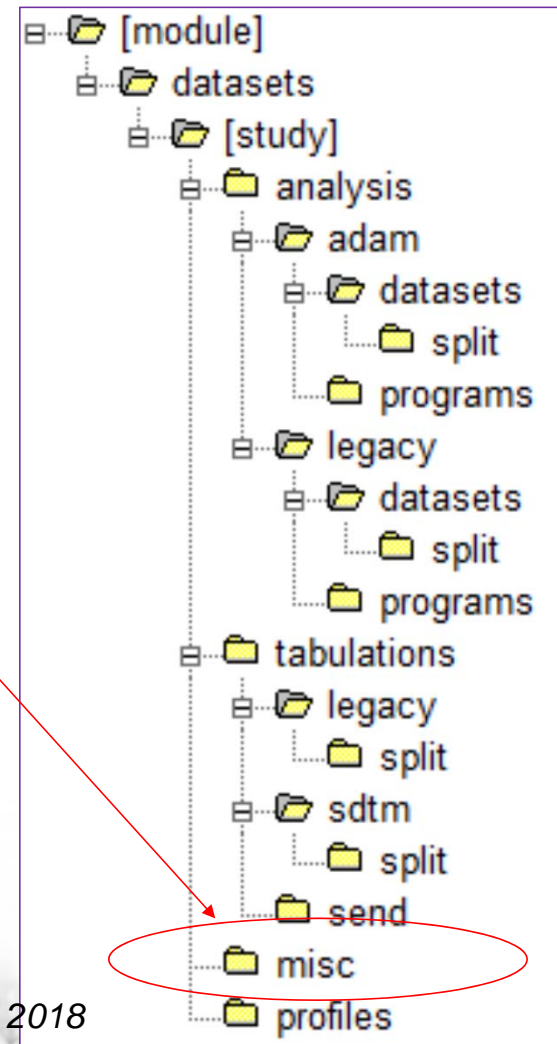
Other Submitted Data

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

- **Submit programs as ASCII text files (*.txt)**
 - Example: **adtte.sas** submitted as **adtte.txt**
 - Note that FDA Reviewers may not use SAS
- **Recommendations**
 - **Submit programs for at least each dataset submitted and all 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

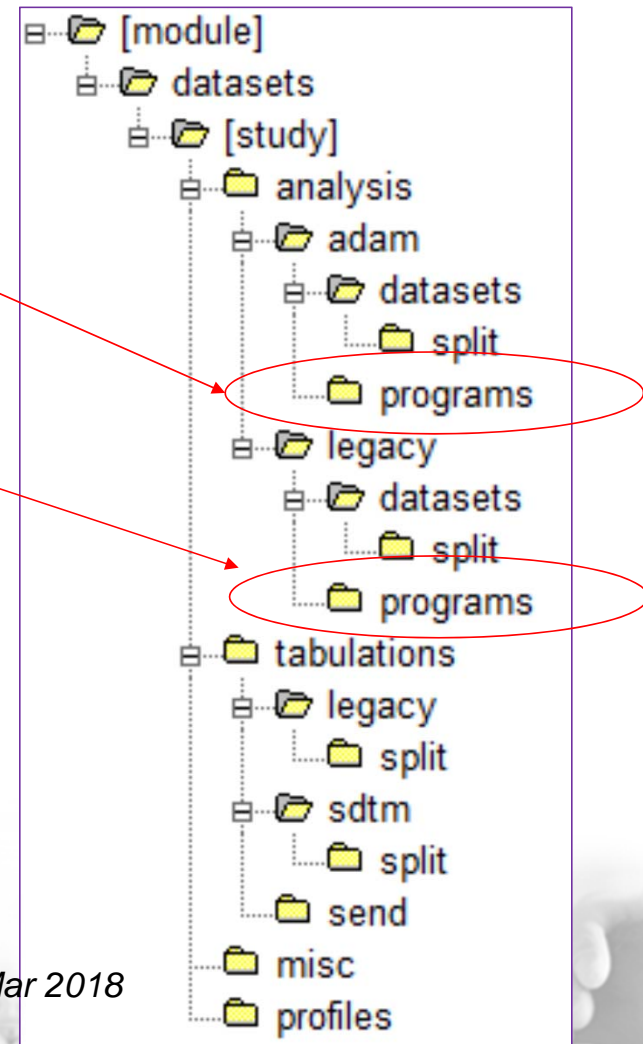
Analysis Programs Location

Place study programs in this subfolder if the study datasets are in the **adam/datasets** folder

OR Place study programs in this subfolder if the study datasets are in the **legacy/datasets** folder

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



References: FDA Study Data Technical Conformance Guide, section 7, Mar 2018
ADaMIG v1.1, Feb 2016

Preparing ADaM and Related Files for Submission

What is Submitted?

- Analysis and other related data
- Analysis programs
- Define files
- ADRG

Data Definition Files

- The data definition (define) file
 - Describes the metadata of the submitted electronic datasets
 - 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
 - Provide a printable define.pdf when define.xml cannot be printed
 - If define.xml v2.0 is used, there is no printing issue
- **Can I submit define.xml v1.0?**
 - Define.xml v1.0 is in FDA Data Standards Catalog, with support ending March, 2018
 - A waiver can be requested for this earlier standard

References: *FDA Study Data Technical Conformance Guide, section 4.1.4.5, Mar 2018*

FDA Data Standards Catalog v4.25.18, Apr 2018

Providing Regulatory Submissions in Electronic Format — Standardized Study Data , Dec 2014

Define Files

Define.XML Content

- An ADaM define.xml typically includes
 - Dataset-level Metadata
 - Variable-level Metadata
 - Parameter Value-level Metadata, when appropriate
 - Results-level Metadata (recommended for critical analyses)
 - Controlled terminology and codes
 - Links to other documents (e.g., SAP, ADRG)

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

Define Files

Define File Location

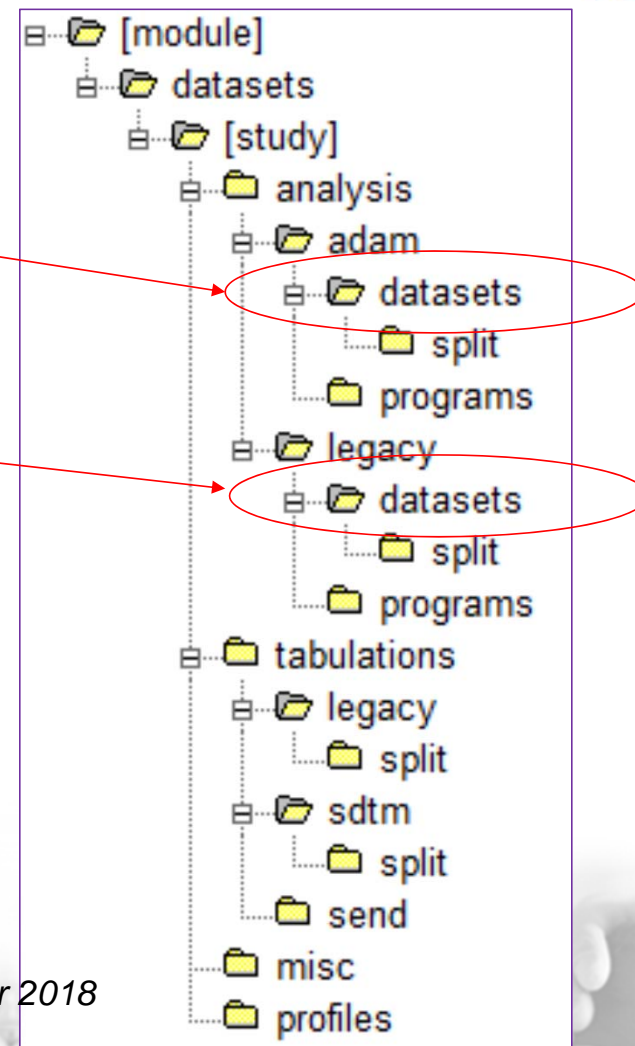
Place define file in this subfolder if the study datasets are in this folder

OR

Place define file in this subfolder if the study datasets are in this folder

Define file sits in the same folder as the data.

Each data folder needs its own define



Define Files

Example adam-> datasets Folder with define files

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

adae.xpt

adef.xpt

adlb.xpt

adsl.xpt

adtte.xpt

adrg.pdf

define.pdf

define.xml

define.xsl

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
 - Recommended in a standards-compliant analysis data submission
 - Provides FDA reviewers with context for analysis datasets and terminology,
 - In addition to what is presented within define.xml
- 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

ADRG

Standard Content

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

ADRG

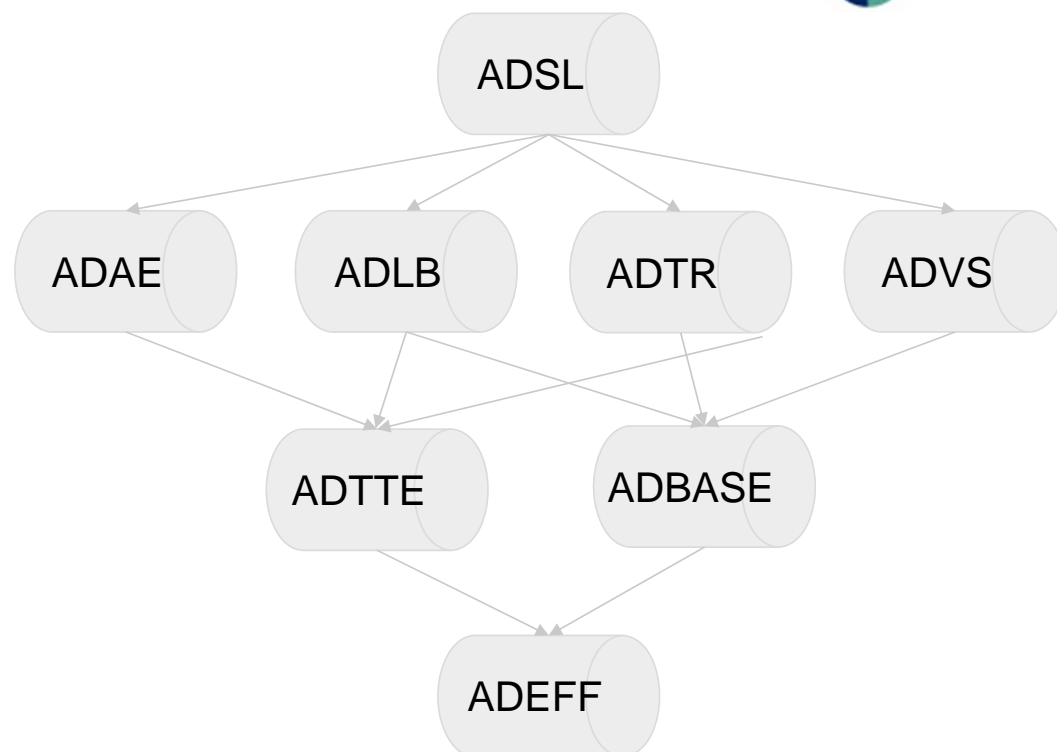
Example Content

- 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**



ADRG

Location

Belongs in the same folder as the datasets and define files

adae.xpt

adef.xpt

adlb.xpt

adsl.xpt

adtte.xpt

adrg.pdf

define.pdf

define.xml

define.xsl

Recap: Preparing ADaM and Related Files for Submission

What is Submitted?

- Analysis and other related data
- Analysis programs
- Define files
- ADRG

Recap

**Datasets
(SAS v5
transport)**

- adae.xpt
- adeff.xpt
- adlb.xpt
- adsl.xpt
- adtte.xpt

ARDG

adrg.pdf

**Define
files**

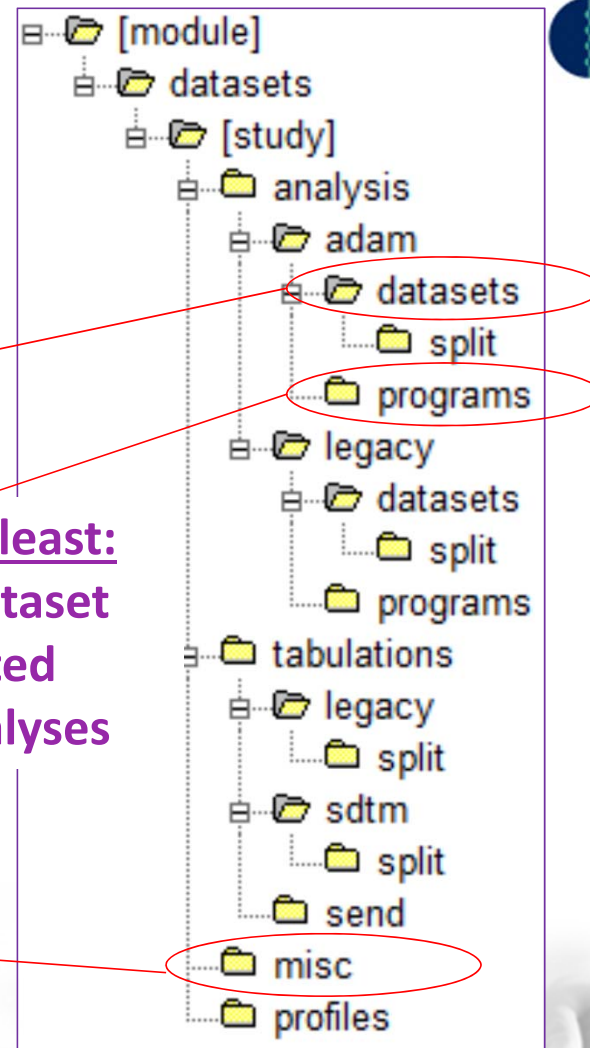
- define.pdf
- define.xml
- define.xsl

Programs for at least:

- Each dataset submitted
- Key analyses

Other data, such as:

- Look-up tables
- Deviations not collected via CRF



Questions?

**Datasets
(SAS v5
transport)**

adae.xpt
adeff.xpt
adlb.xpt
adsl.xpt
adtte.xpt

ARDG

adrg.pdf

**Define
files**

define.pdf
define.xml
define.xsl

Programs for at least:

- Each dataset submitted
- Key analyses

Other data, such as:

- Look-up tables
- Deviations not collected via CRF

