CHANGES IN FDA TECHNICAL CONFORMANCE GUIDE V3.0

13-Sep-2016, Berlin

Marion Friebel

PAREXEL
YOUR JOURNEY. OUR MISSION.℠

DEFINITION

- This Study Data Technical Conformance Guide provides specification, recomendations and general considerations on how to submit standardized study data using FDA – support

- The guide is separated in sections as before:

Section 1: **Introduction** – provides information on regulatory policy and guidance background, purpose, and document control.

Section 2: **Planning and Providing Standardized Study Data** – recommends and provides details on preparing an overall study data standardization plan, a study data reviewer's guide and an analysis data reviewer's guide.

Section 3: **Exchange Format - Electronic Submissions** – presents the specifications, considerations, and recommendations for the file formats currently supported by FDA.

CONTINUE

Section 4: **Study Data Submission Format: Clinical and Nonclinical** – presents general considerations and specifications for sponsors using, for example, the following standards for the submission of study data: Study Data Tabulation Model (SDTM), Analysis Data Model (ADaM), and Standard for Exchange of Nonclinical Data (SEND).

Section 5: **Therapeutic Area Standards** – presents supplemental considerations and specific recommendations when sponsors submit study data using FDA-supported therapeutic area standards (TA).

Section 6: **Terminology** – presents general considerations and specific recommendations when using controlled terminologies/vocabularies for clinical trial data.

Section 7: **Electronic Submission Format** – provides specifications and recommendations on submitting study data using the electronic Common Technical Document (eCTD) format.

Section 8: **Data Validation and Traceability** – provides general recommendations on conformance to standards, data validation rules, data traceability expectations, and legacy data conversion.

CHANGES FROM GUIDE V2.3 TO GUIDE V3.0 (RELEASED MARCH 2016) –UPDATES TO SECTION 2

V2.2

2.2 Study Data Reviewer's Guide

v3.0

- In both versions the text is the same, but for version 2.3 the link doesn't work anymore. Old link refers a new place but not a concrete link is mentioned.

CONTINUE - UPDATE IN SECTION 3.3.2

V2.2

Each dataset should be provided in a single transport file. The maximum size of an individual dataset that FDA can process depends on many factors. Datasets greater than 1 gigabyte (gb) in size should be split into smaller datasets no larger than 1 gb. Sponsors should submit these smaller datasets, in addition to the larger non-split datasets, to better support regulatory reviewers. The split datasets should be placed in a separate sub-directory labeled "split" (See section 7). Clear explanation regarding how these datasets were split needs to be presented within the relevant data reviewer's guide (i.e., SDRG or ADRG).

V3.0

Each dataset should be provided in a single transport file. The maximum size of an individual dataset that FDA can process depends on many factors. Datasets greater than 5 gigabytes (GB) in size should be split into smaller datasets no larger than 5 GB. Sponsors should submit these smaller datasets, in addition to the larger non-split datasets, to better support regulatory reviewers. The split datasets should be placed in a separate sub-directory labeled "split" (See section 7.1). A clear explanation regarding how these datasets were split needs to be presented within the relevant data reviewer's guide (i.e., SDRG or ADRG).

CONTINUE - UPDATES IN SECTION 4.1.

V2.2

4.1.1.2 SDTM General Considerations

The SDTMIG should be followed unless otherwise indicated in this Guide or in the *Standards Catalog*. The conformance criteria listed in the SDTMIG should not be interpreted as the sole determinant of the adequacy of submitted data. If there is uncertainty regarding implementation, the sponsor should discuss application-specific questions with the review division and general standards implementation questions with the specific center resources identified elsewhere in this Guide (See section 1.2). No data should be imputed in SDTM datasets. Data should only be imputed in ADaM datasets (See section 4.1.2.9.2).

V3.0

4.1.2.2 General Considerations

The SDTMIG should be followed unless otherwise indicated in this Guide or in the *Standards Catalog*. The conformance criteria listed in the SDTMIG should not be interpreted as the sole determinant of the adequacy of submitted data. If there is uncertainty regarding implementation, the sponsor should discuss application-specific questions with the review division and general standards implementation questions with the specific center resources identified elsewhere in this Guide (See section 1.2).

Each submitted SDTM dataset should have its contents described with complete metadata in the define.xml file (See section 4.1.4.5) and within the SDRG as appropriate (See section 2.2). No data should be imputed in SDTM datasets. Data should only be imputed in ADaM datasets (See section 4.1.2.9).

Generally, ADaM facilitates FDA review. One of the expected benefits of analysis datasets that conform to ADaM is that they simplify the programming steps necessary for performing an analysis. As noted above, ADaM datasets should be derived from the data contained in the SDTM datasets. There are features built into the ADaM standard that promote traceability from analysis results to ADaM datasets and from ADaM datasets to SDTM datasets. To ensure traceability, all SDTM variables utilized for variable derivations in ADaM should be included in the ADaM datasets when practical. Each analysis dataset that is submitted should be described accordingly with complete metadata in the define.xml file (See section 4.1.4.5).

Generally, ADaM facilitates FDA review. One of the expected benefits of analysis datasets that conform to ADaM is that they simplify the programming steps necessary for performing an analysis. As noted above, ADaM datasets should be derived from the data contained in the SDTM datasets. There are features built into the ADaM standard that promote traceability from analysis results to ADaM datasets and from ADaM datasets to SDTM datasets. To ensure traceability, all SDTM variables utilized for variable derivations in ADaM should be included in the ADaM datasets when practical. Each submitted ADaM dataset should have its contents described with complete metadata in the define.xml file (See section 4.1.4.5) and within the ADRG as appropriate (See section 2.3).

CONTINUE - UPDATES IN SECTION 4.1.

V2.2

4.1.3.2 General Considerations

The SENDIG provides specific domain models, assumptions, conformance and business rules, and examples for preparing standard tabulation datasets that are based on the SDTM. If there is uncertainty regarding SEND implementation, the sponsor should discuss the issue with the review division.

The ideal time to implement SEND is prior to the conduct of the study as it is very important that the results presented in the accompanying study report be traceable back to the original data collected.

4.1.4.5 Data Definition Files for SDTM, SEND, and ADaM

The data definition file describes the metadata of the submitted electronic datasets, and is considered arguably the most important part of the electronic dataset submission for regulatory review. This data definition specification for submitted datasets defines the metadata structures that should be used to describe the datasets and variables. An insufficiently documented data definition file is a common deficiency that reviewers have noted. Consequently, the sponsor needs to provide complete detail in this file, especially for the specifications pertaining to derived variables. In addition, sponsors should also make certain that the code list and origin for each variable are clearly and easily accessible from the data definition file. The version of any external dictionary should be clearly stated both in the data definition file and, where possible, in the updated Trial Summary (TS) domain (i.e., SDTMIG 3.1.2 or greater; SENDIG 3.0 or greater). The internal dataset label should also clearly describe the contents of the dataset. For example, the dataset label for an efficacy dataset might be "Time to Relapse (Efficacy)."

V3.0

4.1.3.2 General Considerations

The SENDIG provides specific domain models, assumptions, conformance and business rules, and examples for preparing standard tabulation datasets that are based on the SDTM. If there is uncertainty regarding SEND implementation, the sponsor should discuss the issue with the review division.

The ideal time to implement SEND is prior to the conduct of the study as it is very important that the results presented in the accompanying study report be traceable back to the original data collected. Each submitted SEND dataset should have its contents be described with complete metadata in the define.xml file (See section 4.1.4.5) and within the SDRG as appropriate (See section 2.2).

4.1.4.5 Data Definition Files for SDTM, SEND, and ADaM

The data definition file describes the metadata of the submitted electronic datasets, and is considered arguably the most important part of the electronic dataset submission for regulatory review. This data definition specification for submitted datasets defines the metadata structures that should be used to describe the datasets, variables, possible values of variables when appropriate, and controlled terminologies and codes. An insufficiently documented data definition file is a common deficiency that reviewers have noted. Consequently, the sponsor needs to provide complete detail in this file, especially for the specifications pertaining to derived variables. In addition, sponsors should also make certain that the code list and origin for each variable are clearly and easily accessible from the data definition file. The version of any external dictionary should be clearly stated both in the data definition file and, where possible, in the updated Trial Summary (TS) domain (i.e., SDTMIG 3.1.2 or greater; SENDIG 3.0 or greater). The internal dataset label should also clearly describe the contents of the dataset. For example, the dataset label for an efficacy dataset might be "Time to Relapse (Efficacy)."

Possible values of variables→ annotation on aCRF

CONTINUE - UPDATES IN SECTION 4.1.

4.1.4.5 Data Definition Files for SDTM, SEND, and ADaM

V2.2

define.pdf should be provided if the define.xml cannot be printed [28]. To confirm that a define.xml is printable within the CDER IT environment, it is recommended that the sponsor submit a test version to cder-edata@fda.hhs.gov prior to application submission. If a define.xml version 2.0 or later version is submitted, then a define.pdf does not need to be included in the submission. The *Standards Catalog* lists the currently supported version(s) of define.xml. Sponsors should include a reference to the style sheet as defined in the specification and place the corresponding style sheet in the same submission folder as the define.xml file.

V3.0

define.pdf should be provided if the define.xml cannot be printed [29]. To confirm that a define.xml is printable within the CDER IT environment, it is recommended that the sponsor submit a test version to cder-edata@fda.hhs.gov prior to application submission. The *Standards Catalog* lists the currently supported version(s) of define.xml. It should be noted that define.xml version 2.0 is the preferred version. Sponsors should include a reference to the style sheet as defined in the specification and place the corresponding style sheet in the same submission folder as the define.xml file.

Define.xml version 2.0 is preferred version

CONTINUE - UPDATES IN SECTION 5.

V2.2

5.1 General

For an SDTM domain associated with a therapeutic area user guide, sponsors should contact the appropriate review division.

V3.0

5.1 General

CDISC Therapeutic Area Standards are comprised of existing data elements, but may introduce new data elements (e.g. domains, variables, terminologies). These data elements are components of current CDISC implementation guides or will be integrated into future implementation guides. CDISC publishes a user guide for each therapeutic area use case which describes the most common data elements for clinical studies (http://www.cdisc.org/therapeutic).

5.2 Supported Therapeutic Areas

Generally, when a data standard is released for public use by the SDO, it is not supported by FDA and is not listed in the FDA Data Standards Catalog. FDA will perform acceptance testing on the standard to confirm its ability to process, review and archive. The CDISC data elements associated with following therapeutic areas are supported by FDA:

5.2.1 Chronic Hepatitis C

5.2.2 Dyslipidemia

CONTINUE - UPDATES IN SECTION 6.1.2.1

V2.2

V3.0

6.1.2.1 Use of the specific controlled term "OTHER"

It is understood that the expansion of controlled terminology may lag behind scientific advancement, and that sometimes there may not be a relevant term within a controlled terminology's value set to describe a clinical trial event, finding, or observation. However, it is not recommended to map a collected value to "OTHER" when there is a controlled term available to match the collected value – even when the terminology allows for Sponsor expansion. Each unique value in a --TERM field mapped to a --DECODE value of "OTHER" should have a clear rationale outlined in the Study Data Reviewer's Guide (clinical or non-clinical).

Not recommended to use „OTHER" if controlled term is available, if the word „OTHER" in ~TERM variables then they are not unique anymore

CONTINUE - UPDATES IN SECTION 8.3.1

V2.2

studies are prospectively designed to collect data using a standardized CRF, e.g., CDASH.

As noted in section 1.1, the submission of standardized study data will be required according to the timetable specified in the eStudy Data guidance. During the transition period to required study data standards, FDA recognizes that some study data (i.e., legacy data) submissions may not conform to FDA-supported study data standards and may need to be converted.

V3.0

studies are prospectively designed to collect data using a standardized CRF, e.g., CDASH. Traceability can be further enhanced when a flow diagram is submitted showing how data move from collection through preparation and submission to the Agency.

As noted in section 1.1, the submission of standardized study data will be required according to the timetable specified in the eStudy Data guidance. During the transition period to required study data standards, FDA recognizes that some study data (i.e., legacy data) submissions may not conform to FDA-supported study data standards and may need to be converted.

Flow diagram is recommended to show traceability

SUMMERY OF CHANGES FROM VERSION 2.2 TO 3.0

- BLANKCRF.PDF RENAMED TO ACRF.PDF
- DEFINE.XML VERSION 2.0 IS NOW THE PREFERRED VERSION
- DATASETS GREATER THAN 5 GB SHOULD BE SPLIT (BEFORE THE LIMIT WAS 2 GB)
- IT IS NOT RECOMMENDED TO USE „OTHER“ IN VARIABLES AND CT
- A FLOW DIAGRAM IS RECOMMENDED TO SHOW TRACEABILITY
- TAUGS ARE COMPRISED OF EXISTING DATA ELEMENTS , NEW ELEMENTS WILL GO INTO NEW GUIDES, NEW TAUGS
- TAUGS RELEASED FOR PUBLIC BUT NOT ALL SUPPORTED BY FDA
- FDA OFFERS HELP TO SMOOTH THE PROCESS (TEST SUBMISSIONS , TOOLS FOR VALIDATION)

THANK YOU

