INSTRUCTIONS TO SELF-CHECK WORKSHEET FOR STUDY DATA PREPARATION

(The field numbers below correspond to the numbered boxes in each section of the Self-Check Worksheet for Study Data Preparation)

NOTE: The Agency performed a study data validation assessment utilizing the validation rules as specified in FDA's Technical Rejection Criteria for Study Data.[1] The checklist is for industry to validate the technical conformance of clinical and nonclinical study datasets included in a submission with study data.

The Self-Check Worksheet for Study Data Preparation is designed to check the validation criteria for newly submitted Study Data. Section 1 should be completed only once for a submission with study data. Section 2 through Section 5 can be repeated for each study included in the submission.

Technical specifications associated with this guidance are provided in FDA's Technical Rejection Criteria for Study Data[2], Study Data Technical Conformance Guide[3], and other stand-alone documents that are updated periodically, which can be found on FDA's Study Data website[4]:
- Data Standards Catalog
- FDA Business Rules
- FDA Validator Rules

The following four high-level study data validation rules per FDA's Technical Rejection Criteria for Study Data[5] are checked:

Validation Error	Description
Validation Error Number 1734:	A dataset named ts.xpt with information on study start date must be present for each study in module 4, sections 4.2.3.1, 4.2.3.2, 4.2.3.4, and in module 5, sections 5.3.1.1, 5.3.1.2, 5.3.3.1, 5.3.3.2, 5.3.3.3, 5.3.3.4, 5.3.4, 5.3.5.1, 5.3.5.2
Validation Error Number 1735:	The correct STF file-tags must be used for all standardized datasets and corresponding define.xml files in module 4, sections 4.2.3.1, 4.2.3.2, 4.2.3.4, and in module 5, sections 5.3.1.1, 5.3.1.2, 5.3.3.1, 5.3.3.2, 5.3.3.3, 5.3.3.4, 5.3.4, 5.3.5.1, 5.3.5.2
Validation Error Number 1736:	For SEND data, a DM dataset and define.xml must be submitted in module 4, sections 4.2.3.1, 4.2.3.2, 4.2.3.4 For SDTM data, a DM dataset and define.xml must be submitted module 5, sections 5.3.1.1, 5.3.1.2, 5.3.3.1, 5.3.3.2, 5.3.3.3, 5.3.3.4, 5.3.4, 5.3.5.1, 5.3.5.2 For ADaM data, an ADSL dataset and define.xml must be submitted in module 5, sections 5.3.1.1, 5.3.1.2, 5.3.3.1, 5.3.3.2, 5.3.3.3, 5.3.3.4, 5.3.4, 5.3.5.1, 5.3.5.2
Validation Error Number 1789:	A file has been submitted in a study section without providing an STF file. STFs are not required for 4.3 Literature references, 5.2 Tabular listings, 5.4 Literature references and 5.3.6 Postmarketing reports

---

[1] Technical Rejection Criteria for Study Data

[2] ibid

[3] Study Data Technical Conformance Guide

[4] Resources related to study data, including guidance documents, technical guides, business rules, and position statements, can be found on the FDA Study Data website.

[5] See footnote 1.

Version: 11/30/2019

INSTRUCTIONS TO SELF-CHECK WORKSHEET FOR STUDY DATA PREPARATION

*Required field

Section 1: Application & Submission Information
Reference: 21CFR 314.3(b)

Field 1a. FDA Center*: Identify the appropriate Center by which the application will be reviewed.
Field 1b. Application Type*: Identify the appropriate application type.
Field 1c. Application Number: Provide the six-digit application number.
Field 1d. eCTD Sequence Number: Provide the four-digit eCTD sequence number.
Field 1e. eCTD Submission Type: Provide an appropriate submission type.
Field 1f. eCTD Submission Sub Type: Identify submission sub-types.

SECTION 2 THROUGH SECTION 5 CAN BE REPEATED FOR EACH STUDY INCLUDED IN THE SUBMISSION

Section 2: Study Information
Reference: N/A

Field 2a. Study ID*: Identify the Study ID for the study being submitted. Study ID is the unique identifier across application documents. Therefore, the Study ID must be consistent across all the files being submitted for the same study, i.e. STF File, ts.xpt, dm.xpt, etc. For the purposes of this worksheet, **Fields 2a, 3d, 4f/4h and 4l** should include the same study ID.
Field 2b. Is This the First Time Study Data is Being Submitted for This Study as Part of This Application?*: Indicate whether the study data is being submitted for the first time. The Self-Check Worksheet is designed for newly submitted study data, i.e. studies for which no files have been previously submitted.
Field 2c. Title of the Study: Provide title of the study that is being submitted.
Field 2d. Study Section - eCTD Heading (Example: m4-2-1-1)*: Identify study section/eCTD Heading under which the Study Data is being submitted (Example m4-2-1-1).
Field 2e. Module*: Identify an appropriate module that the Study Data is associated with. If nonclinical study data is being submitted, then select "Nonclinical (m4)". If clinical Study Data is being submitted, then select "Clinical (m5)".
Field 2f. Study Dataset Type(s)*: Identify the Study Data associated with the Study ID in **Field 2a** as tabulation and/or analysis, or neither tabulation or analysis, i.e. "Other." If you are submitting tabulation data, select "Tabulation." If you are submitting analysis data, select "Analysis." For other types of data, such as Listings datasets (when tabulation or analysis data is not being submitted), or text-based documents with no tabulated data or line listings included, select "Other."

Section 3: STF File Information
Reference: Validation Error Number 1789[6]

Field 3a. Are Files Included in a Study Section? (Not Applicable to Sections 4.3, 5.2, 5.3.6, and 5.4)*: If the Study Data for the study identified in **Field 2a** contains files that are not included in sections 4.3, 5.2, 5.3.6, or 5.4, select "Yes". If the Study Data contains files in sections 4.3, 5.2, 5.3.6, or 5.4 *only*, select "No".
Field 3b. Is STF File Included?*: All study files submitted in Study Sections other than the ones listed in **Field 3a** should have an associated STF file. If the Study Data identified in **Field 2a** contains STF File, select "Yes". If the Study Data identified in **Field 2a** does not contain STF File, select "No".
Field 3c. Does STF File Reference all Associated Study Files?*: Identify whether all files associated with the study are referenced in an STF file.
Field 3d. Study ID in STF File*: Identify Study ID in the STF File of the Study Data identified in **Field 2a**. Study ID in STF File in **Field 3d** should be identical to the Study ID identified in **Field 2a**.
Field 3e. Does the Study ID in the STF File Match Field 2a?: Identify Study ID provided in **Field 2a** and **Field3d** match. Study ID should be consistent across all the files being submitted for the same study. This field will be preselected based on previous inputs.
Field 3f. Are XPT Datasets (other than the TS File) Included?*: Identify if XPT datasets (other than the TS File) are included among study files. The inclusion of XPT datasets helps determine TS File requirements in **Section 4** for submissions with clinical or nonclinical study data.

---

[6] See footnote 1.

Field 3g. If the Study is Nonclinical (m4) , are Any Study Files Tagged as "pre-clinical-study-report," "legacy-clinical-study-report," or "study-report-body"?: If the study is submitted to CDER, and is indicated as nonclinical (m4) in **Field 2e**, identify if the file tags "pre-clinical-study-report," "legacy-clinical-study-report," or "study-report-body" are used for one or more study reports in the STF File.

Section 4: TS File Information
Reference: Validation Error Number 1734[7]

Field 4a. Is your Study Start Date (Commercial IND): **Prior to or on 17-Dec-2017?** or **After 17-Dec-2017?** Identify date the clinical or non-clinical study officially began.[8]

Field 4b. Is your Study Start Date (NDA, BLA, or ANDA): Prior to or on 17-Dec-2016? or **After 17-Dec-2016?** Identify date the clinical or non-clinical study officially began.[9]

Field 4c. Is TS file required?* Refer to current Technical Rejection Criteria (TRC).[10] If files are included in a Study Section in module 4, sections 4.2.3.1, 4.2.3.2, 4.2.3.4, and in module 5, sections 5.3.1.1, 5.3.1.2, 5.3.3.1, 5.3.3.2, 5.3.3.3, 5.3.3.4, 5.3.4, 5.3.5.1, 5.3.5.2, then TS File is required.

Field 4d. If TS File required, is it included?*: Identify whether the Study Data identified in **Field 3a** includes TS File.

Field 4e If TS File is Required, What Type of TS File is Required? Full TS or **Simplified TS.** Refer to *Table 1: eCTD Technical Rejection Criteria for Study Data Expectations* in the current Technical Rejection Criteria (TRC).[11]

Select **"Full TS"** (1) if Study Start Date for a clinical study is after December 17, 2016 for NDA, ANDA, or BLA submitted to CDER or CBER; (2) if Study Start Date for a nonclinical study is after December 17, 2016 for NDA, ANDA, or BLA submitted to CDER; (3) if Study Start Date for a nonclinical study is after December 17, 2017 for Commercial IND submitted to CDER.

Select "**Simplified TS**" (1) if Study Start Date for a clinical study containing an XPT dataset (other than the ts.xpt) is on or prior to December 17, 2016 for NDA, ANDA, or BLA submitted to CDER or CBER; (2) if Study Start Date for a nonclinical study containing an XPT dataset (other than the ts.xpt) or using the file tags "pre-clinical-study-report," "legacy-clinical-study-report," or "study-report-body" is on or prior to December 17, 2016 for NDA, ANDA, or BLA submitted to CDER; (3) if Study Start Date for a nonclinical study containing an XPT dataset (other than the ts.xpt) or using the file tags "pre-clinical-study-report," "legacy-clinical-study-report," or "study-report-body" is on or prior to December 17, 2017 for Commercial IND submitted to CDER.

Full TS File

Field 4f. Study ID in TS File*: Identify the Study ID in Full TS File.

Field 4g. Does Study ID in STF (Field 3d) & TS Files Match?*: Identify whether the Study ID in Full TS File in **Field 4f.** is identical to Study ID in **Field 3d**. This field will be preselected based on previous inputs.

Field 4h. If Study ID does Not Match, What is the Value of SPREFID in TS File?: Identify the SPREFID value in the TS File. For studies where alignment of the study identifier across TS and STF is not feasible, the value for [study-id] used in the STF should be included in TS using the parameter SPREFID. SPREFID Can be used to reconcile study identifiers where necessary for SEND or SDTM studies.

Field 4i. Does Study ID in STF (Field 3d) and SPREFID Match?: Identify whether the Study ID in the TS File in **Field 4h.** is identical to Study ID in **Field 3d**. This field will be preselected based on previous inputs.

Field 4j. Study Start Date in TS File: Identify Study Start Date as it appears in TS File of study identified in **Field 2a**. If the Study Start Date is not available, leave the field blank.

Field 4k. If Study Start Date Exists, Is it in Valid Format (yyyy-mm-dd)?: If the Study Start Date is provided in **Field 4j**, identify whether the Study Start Date is valid, i.e. in the valid format: yyyy-mm-dd. Please note that the Study Start Date (SSD) should follow the ISO 8601 standard that provides, at a minimum, the year, month, and day.

---

[7] See footnote 1.

[8] See footnote 3.

[9] See footnote 3.

[10] See footnote 1.

[11] ibid.

Field 4l. Study ID in TS File*: Identify the Study ID in Simplified TS File.
Field 4m. Does Study ID in STF (Field 3d) & TS Files Match?*: Identify whether the Study ID in the TS File in **Field 4l.** is identical to Study ID in **Field 3d**. This field will be preselected based on previous inputs.
Field 4n. Is there a Value in TSVALNF?: Identify if TSVALNF value is included in Simplified TS File.
Field 4o. Is the Value in TSVALNF "NA"?: Identify if the TSVALNF value is "NA".
Field 4p. Study Start Date in TS File: Identify the Study Start Date in Simplified TS File.
Field 4q. If Study Start Date Exists, Is it in Valid Format (yyyy-mm-dd)?: If the Study Start Date is provided in **Field 4p**, identify whether the Study Start Date is valid, i.e. in the valid format: yyyy-mm-dd. Please note that the Study Start Date (SSD) should follow the ISO 8601 standard that provides, at a minimum, the year, month, and day.

Section 5: Standardized Datasets (SEND, SDTM, ADaM)
Reference: Validation Error Numbers 1735 & 1736[12]

Field 5a. Are Standardized Datasets Required?*: Identify whether standardized datasets are required or not required. Standardized datasets are required: 1) if Study Start Date is after December 17, 2016 for NDA, ANDA, or BLA submitted to CDER or CBER; 2) if Study Start Date is after December 17, 2017 for Commercial IND submitted to CDER or CBER. **Fields 5b-5e** are applicable to non-clinical tabulation datasets (SEND), **Fields 5f-5i** are applicable to clinical tabulation datasets (SDTM), and **Fields 5j-5m** are applicable to clinical analysis datasets (ADaM).

If you answered "No" in **Field 3f** and "Yes" in **Field 3g**, standardized datasets are not required in this submission and Validation Rules 1735 and 1736 do not apply. For more details on standardized dataset requirements, please refer to the Study Data Technical Conformance Guide.[13]

Study Start Date	Application Type	Standardized Datasets Required?
Prior to or on 17-Dec-16	NDA, BLA, or ANDA	Not Required
After 17-Dec-16	NDA, BLA, or ANDA	Required
Prior to or on 17-Dec-17	Commercial IND	Not Required
After 17-Dec-17	Commercial IND	Required

Nonclinical (m4): Tabulation (SEND datasets)

Field 5b. Is DM File Included?*: Identify whether a DM dataset is included. DM dataset is required for submissions in Study Section module 4, section 4.2.3.1, 4.2.3.2, 4.2.3.4.
Field 5c. Is Define File Included?*: Identify whether a Define File is included. Define File is required for submissions in Study Section module 4, section 4.2.3.1, 4.2.3.2, 4.2.3.4.
Field 5d. Are the STF File-Tags for the SEND Datasets "data-tabulation-dataset-send"?*: Identify that all the datasets (XPT files) use the correct STF file-tags in module 4, section 4.2.3.1, 4.2.3.2, 4.2.3.4. Valid file tag for XPT files for module 4 is "data-tabulation-dataset-send".
Field 5e. Is the STF File-Tag for Define File "data-tabulation-data-definition"?*: Define File is required for submissions in Study Section in module 4, section 4.2.3.1, 4.2.3.2, 4.2.3.4 of the tabulation dataset. Valid file tag for Define File for module 4 is "data-tabulation-dataset-definition".

Clinical (m5): Tabulation (SDTM datasets)

---

[12] See footnote 1.

[13] See footnote 3.

Field 5f. Is DM File Included?*: Identify whether a DM dataset is included. DM dataset is required for submissions in Study Section module 5, sections 5.3.1.1, 5.3.1.2, 5.3.3.1, 5.3.3.2, 5.3.3.3, 5.3.3.4, 5.3.4, 5.3.5.1, 5.3.5.2 of the tabulation dataset.

Field 5g. Is Define File Included?*: Identify whether a Define File is included. Define File is required for submissions in Study Section module 5, sections 5.3.1.1, 5.3.1.2, 5.3.3.1, 5.3.3.2, 5.3.3.3, 5.3.3.4, 5.3.4, 5.3.5.1, 5.3.5.2 of the tabulation dataset.

Field 5h. Are the STF File-Tags for the SDTM Datasets "data-tabulation-dataset-sdtm"?*: Identify that all the datasets (XPT files) use the correct STF file-tags in module 5, sections 5.3.1.1, 5.3.1.2, 5.3.3.1, 5.3.3.2, 5.3.3.3, 5.3.3.4, 5.3.4, 5.3.5.1, 5.3.5.2. Valid file tag for XPT files for module 5 tabulation dataset is "data-tabulation-dataset-sdtm"

Field 5i. Is the STF File-tag for the Define File "data-tabulation-data-definition?*: Define File is required for submissions in Study Section in module 5, sections 5.3.1.1, 5.3.1.2, 5.3.3.1, 5.3.3.2, 5.3.3.3, 5.3.3.4, 5.3.4, 5.3.5.1, 5.3.5.2 of the tabulation dataset. Valid file tag for corresponding define.xml file for module 5 tabulation dataset is "data-tabulation-data-definition".

Clinical (m5): Analysis (ADaM datasets)

Field 5j. Is ADSL File Included?*: Identify whether a ADSL dataset is included. ADSL dataset is required for submissions in Study Section module 5, sections 5.3.1.1, 5.3.1.2, 5.3.3.1, 5.3.3.2, 5.3.3.3, 5.3.3.4, 5.3.4, 5.3.5.1, 5.3.5.2 of the analysis dataset.

Field 5k. Is Define File Included?*: Identify whether a Define File is included. Define File is required for submissions in Study Section module 5, sections 5.3.1.1, 5.3.1.2, 5.3.3.1, 5.3.3.2, 5.3.3.3, 5.3.3.4, 5.3.4, 5.3.5.1, 5.3.5.2 of the analysis dataset.

Field 5l. Are the STF File-Tags for the ADaM Datasets "analysis-dataset-adam"?*: Identify that all the datasets (XPT files) use the correct STF file-tags in module 5, sections 5.3.1.1, 5.3.1.2, 5.3.3.1, 5.3.3.2, 5.3.3.3, 5.3.3.4, 5.3.4, 5.3.5.1, 5.3.5.2. Valid file tag for XPT files for module 5 analysis dataset is "analysis-dataset-adam"

Field 5m. Is the STF File-Tag for the Define File "analysis-data-definition"?*: Define File is required for submissions in Study Section in module 5, sections 5.3.1.1, 5.3.1.2, 5.3.3.1, 5.3.3.2, 5.3.3.3, 5.3.3.4, 5.3.4, 5.3.5.1, 5.3.5.2 of the analysis dataset. Valid file tag for corresponding Define File for module 5 analysis dataset is "analysis-data-definition".

If you have further questions or need clarification, please contact edata@fda.hhs.gov.