Submitting Nonclinical Datasets for Evaluation of Rodent Carcinogenicity Studies of Pharmaceuticals

Guidance for Industry

Technical Specifications Document

For questions regarding this technical specifications document, contact CDER at cder-edata@fda.hhs.gov.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)

[May 2021]
Technical Specifications Document

Revision History

Date	Version	Summary of Revisions
May 2021	1.0	Initial Version

Table of Contents

Submitting Nonclinical Datasets for Evaluation of Rodent Carcinogenicity Studies of Pharmaceuticals

Guidance for Industry
Technical Specifications Document[1]

This guidance represents the current thinking of the Food and Drug Administration (FDA or Agency) on this topic. It does not establish any rights for any person and is not binding on FDA or the public. You can use an alternative approach if it satisfies the requirements of the applicable statutes and regulations. To discuss an alternative approach, contact the FDA office responsible for this guidance as listed on the title page.

1.0 Introduction

This document provides detailed information and specifications for the content of datasets that should be submitted as part of the sponsor's or applicant's application for drugs that are assessed for their evaluation of rodent carcinogenicity studies of pharmaceuticals in nonclinical studies. These specifications also provide an opportunity for dialogue between the sponsor or applicant and reviewers to discuss issues with trial design or study conduct that may affect the content of the analysis datasets. These specifications were built to support the data standards and processes described in the FDA Study Data Technical Conformance Guide (TCG) and guidance for industry *Providing Regulatory Submissions in Electronic Format — Standardized Study Data* (October 2020).[2]

For questions about a particular data standard implementation, send an email to cder-edata@fda.hhs.gov. For more general recommendations on the use and submission of standardized study data, the sponsor or applicant should refer to the TCG.

In general, FDA's guidance documents do not establish legally enforceable responsibilities. Instead, guidances describe the Agency's current thinking on a topic and should be viewed only as recommendations, unless specific regulatory or statutory requirements are cited. The use of

[1] This guidance was prepared by the Office of Biostatistics in the Center for Drug Evaluation and Research at the Food and Drug Administration. You may submit comments on this guidance at any time. Submit comments to Docket No. FDA-2018-D-1216 (available at https://www.regulations.gov/docket?D=FDA-2018-D-1216) (see the instructions for submitting comments in the docket).

[2] The FDA Study Data Technical Conformance Guide is available at https://www.fda.gov/industry/fda-resources-data-standards/study-data-standards-resources. We update guidances periodically. For the most recent version of a guidance, check the FDA guidance web page at https://www.fda.gov/regulatory-information/search-fda-guidance-documents.

the word *should* in Agency guidances means that something is suggested or recommended, but not required.

2.0 Overview of the SEND Datasets

The Standard for Exchange of Nonclinical Data (SEND) provides the organization, structure, and format of standard nonclinical (animal toxicology studies) tabulation datasets for regulatory submission. The SEND Implementation Guide (SENDIG) provides specific domain models, assumptions, and examples for preparing standard tabulation datasets. It also guides the organization, structure, and format of standard nonclinical tabulation datasets for interchange between organizations and for submission to regulatory authorities. SENDIG v3.0 supports single-dose general toxicology, repeat-dose general toxicology, and carcinogenicity studies. SENDIG v3.1 additionally supports respiratory and cardiovascular safety pharmacology studies.

According to the guidance for industry *Providing Regulatory Submissions in Electronic Format — Standardized Study Data* (October 2020), the version of SENDIG that the sponsors follow should be based on the nonclinical study initiation date. The ideal time to implement SEND is before conducting the study because it is very important that the results presented in the accompanying study report be traceable back to the original data collected (TCG section 4). Each submitted SEND dataset should have its contents described with complete metadata in the define.xml file (TCG section 4) and in the nonclinical Study Data Reviewers Guide (nSDRG) as appropriate (TCG section 2).

3.0 SEND Data Validation

FDA expects the sponsor to appropriately verify that the data in the SEND datasets are an accurate representation of the data in final reports.

Sponsors should either correct any discrepancies between the SEND datasets and the Clinical Data Interchange Standards Consortium (CDISC) standards or the FDA Business Rules[3], or explain meaningful discrepancies in the nSDRG. Additional information about conformance to the CDISC standards and FDA Business Rules and about establishing consistency and traceability between the study data and the study report should also be provided in the nSDRG (TCG section 8). This information is important for the reviewer to understand the context in which the SEND data are provided. If the reviewer is unable to trace study data from data collected in a study to the analysis of the overall study data, then the SEND data may be of limited utility. The sponsor should ensure the traceability of the SEND data to maximize their utility.

[3] See www.CDISC.org. The FDA Study Data Standards Resources web page provides links to the currently available FDA Business Rules. See https://www.fda.gov/industry/fda-resources-data-standards/study-data-standards-resources.

For every submission, methods should be in place to check validity and accuracy of the SEND datasets, and the results from that verification are expected to be available in the event of Agency request.

4.0 Dataset Specifications for FDA Standard Format of Tumor Data (tumor.xpt)

Carcinogenicity studies should include an electronic dataset of tumor findings to allow for a complete review. Sponsors should continue to include the tumor.xpt and associated define.pdf files regardless of whether the study is submitted with the SEND dataset. When both tumor.xpt and SEND are submitted, the sponsor should ensure that the data are consistent and traceable between the tumor.xpt file and the SEND datasets, and consistent with the information specified in FDA Business Rules. Any information needed to establish traceability between the two datasets should be presented in the nSDRG (TCG section 4).

The Tumor Findings dataset (tf.xpt) is necessary if the SEND datasets are the basis for creation of the tumor.xpt dataset. If sponsors choose not to submit the Tumor Findings dataset (tf.xpt) with the SEND submission, the algorithm used to calculate "Time in days to detection of tumor" should be included in the nSDRG (TCG section 4).

For rodent carcinogenicity studies submitted in module 4.3.2.4 of the electronic Common Technical Document, the tumor.xpt file and its associated define.pdf should be placed in analysis/legacy/dataset subfolder under the study datasets folder (TCG section 7).

4.1 Description of Tumor.xpt – Appendix A

CDER statisticians perform analyses on the tumor data from each rodent carcinogenicity study, and the tumor data should be provided as an electronic analysis dataset. The tumor.xpt (see Appendix A) specifies the recommended data elements to be included in the analysis dataset. The dataset containing this information should be named tumor.xpt to aid in identification.

The tumor.xpt file is a one-record-per-tumor finding per tissue/organ per subject dataset that contains tumor-level variables such as tumor detection date, tumor name, tumor code, tumor malignancy status, cause of death; tissue/organ-level variables such as tissue/organ microscopic examination status, organ name, organ code; and subject-level variables such as death or sacrifice date, death or sacrifice status, animal microscopic examination status, dosing group. This dataset is intended to provide data for primary tumors for statistical analysis. It should be prepared as described in the TCG (section 4). All variables provided in the tumor.xpt should be traceable to the submitted tabulations.

For the tumor.xpt file, sponsors should include a data definition table to describe the format, the content, and the definitions of the variables in the dataset. The data definition table should be provided as a single PDF file named define.pdf and placed in the appropriate study, specific analysis type, or integrated summary folder in the datasets folder.

4.2 Example of Reduced Tumor Data – Appendix B

To illustrate the recommended approach to encode tumor data, a reduced example of tumor.xpt is provided in Appendix B. This reduced example contains only a few tumors per tissue/organ for one subject for illustration purposes.

4.3 Business Rules for Datasets of Carcinogenicity Study – Appendix C

Fourteen business rules (see Appendix C) were developed to support regulatory review and analysis of study data for carcinogenicity studies using tumor.xpt and SEND datasets. These business rules are part of the larger number of business requirements for regulatory review that constitute the FDA Business Rules. According to the TCG (section 8), all FDA Business Rules should be followed if applicable, and the information in SEND datasets should be consistent with those in tumor.xpt. Sponsors should explain any inconsistency in the nSDRG.

5.0 Mapping Procedures From SEND to tumor.xpt

The following domains should be used to create a tumor.xpt file:

- DM (Demographics)
- DS (Disposition)
- EX (Exposure)
- MI (Microscopic Findings)
- TX (Trial Sets)
- TF (Tumor Findings) *

* TF domain is necessary if the SEND datasets are the basis for creation of the tumor.xpt dataset (SENDIGv3.1). If sponsors choose not to submit TF domain with the SEND submission, the algorithm used to calculate "Time in days to detection of tumor" should be included in the nSDRG (TCG section 4).

5.1 Mapping From SEND to tumor.xpt – Appendix D

The table in Appendix D describes the variables in the tumor.xpt and the SEND variables from which they are created.

5.2 Mapping From DSDECOD to DTHSACST

The tumor.xpt dataset's DTHSACST (death or sacrifice status) variable may be mapped from the DS domain's DSDECOD (standardized disposition term) variable. The following table gives the DSDECOD values and their DTHSACST equivalent. The DSDECOD values of MISSING, REMOVED FROM STUDY ALIVE, RECOVERY SACRIFICE, and NON-MORIBUND SACRIFICE do not map to the DTHSACST variable of the tumor.xpt. Subjects with these dispositions would not be relevant for tumor analysis.

DSDECOD	**DTHSACST**	
Value	**Value**	**Meaning**
ACCIDENTAL DEATH	4	Accidental death
FOUND DEAD	1	Natural death or moribund sacrifice
MISSING	*	*
MORIBUND SACRIFICE	1	Natural death or moribund sacrifice
INTERIM SACRIFICE	3	Planned intermittent sacrifice
RECOVERY SACRIFICE	*	*
REMOVED FROM STUDY ALIVE	*	*
NON-MORIBUND SACRIFICE	*	*

* There are no mappings to the DTHSACST variable for these DSDECOD values. Subjects with these dispositions would not be relevant for tumor analysis

6.0 Appendix

Appendix A: FDA Standard Format for Tumor Data (tumor.xpt)

Variable Name	Variable Label	Type	Code	Comments
STUDYNUM	Study number	Char		3
ANIMLNUM	Animal number	Char		1,3
SPECIES	Animal Species	Char	M = Mouse; R = Rat	
SEX	Sex	Char	M = male F = female	

DOSEGP	Dose group	Num	0 = # milligram (mg)/kilogram (kg)/day (Group 1)
			1 = # mg/kg/day (Group 2)
			2 = # mg/kg/day (Group 3)
			3 = # mg/kg/day (Group 4)	

DTHSACTM	Time in days/weeks to death or sacrifice	Num		
DTHSACST	Death or sacrifice status	Num	1 = Natural death or moribund sacrifice 2 = Terminal sacrifice 3 = Planned intermittent sacrifice 4 = Accidental death	
ANIMLEXM	Animal microscopic examination code	Num	0 = No tissues were examined 1 = At least one tissue was examined	
TUMORCOD	Tumor type code	Num		3,4
TUMORNAM	Tumor name	Char		3,4
ORGANCOD	Organ/tissue code	Char		3,5
ORGANNAM	Organ/tissue name	Char	LN = Lymph node GL = Gland NONGL = Nonglandular SALIV GL, MANDIB = Mandibular salivary gland HEMATO=Hematopoietic	3,5
DETECTTM	Time in days/weeks of detection of tumor	Num		
MALIGNST	Malignancy status	Num	1 = Malignant 2 = Benign 3 = Undetermined	4
DEATHCAU	Cause of death	Num	1 = Tumor caused death 2 = Tumor did not cause death 3 = Undetermined	4
ORGANEXM	Organ/tissue microscopic examination code	Num	1 = Organ/tissue was examined and was usable 2 = Organ/tissue was examined but was not usable (e.g., autolyzed tissue or tissue not found in section examined) 3 = Organ/tissue was not examined	

1 Each animal in the study should have at least one record even if it does not have a tumor.
2 Additional variables, as appropriate, can be added to the bottom of this dataset.
3 STUDYNUM and ANIMLNUM limited to no more than 12 characters; ORGANCOD and TUMORCOD limited to no more than 8 characters; ORGANNAM and TUMOR should be as concise as possible.
4 A missing value should be given for the variable MALIGNST, DEATHCAU, TUMORCOD, and TUMORNAM when the organ is unusable or not examined.
5 Do not include a record for an organ that was useable and no tumor was found on examination. A record should be included for organs with a tumor, organs found unusable, and organs not examined.

Appendix B: Example of Reduced FDA Standard Format for Tumor Data (tumor.xpt)

STUDYNUM	ANIMLNUM	SPECIES	SEX	DOSEGP	DTHSACTM	DTHSACST	ANIMLEXM	TUMORCOD	TUMORNAM	ORGANCOD	ORGANNAM	DETECTTM	MALIGNST	DEATHCAU	ORGANEXM
xxxxxxx	1	R	M	1	550	1	1			BODA	body cavity, abdominal				3
xxxxxxx	1	R	M	1	550	1	1			BODN	body cavity, nasal				3
xxxxxxx	1	R	M	1	550	1	1			BODO	body cavity, oral				3
xxxxxxx	1	R	M	1	550	1	1			BODT	body cavity, thoracic				3
xxxxxxx	1	R	M	1	550	1	1	COAR	cortical adenoma	GLAD	gland, adrenal	550	2	2	1
xxxxxxx	1	R	M	1	550	1	1	ADEN	adenoma	GLPI	gland, pituitary	550	2	1	1
xxxxxxx	1	R	M	1	550	1	1			HEMO	hemolymphoreticular tissue				1
xxxxxxx	1	R	M	1	550	1	1	SEBE	seminoma, benign	TEST	testis	550	2	2	1
xxxxxxx	2	R	M	1	498	1	1			BODA	body cavity, abdominal				3
xxxxxxx	2	R	M	1	498	1	1			BODN	body cavity, nasal				3
xxxxxxx	2	R	M	1	498	1	1			BODO	body cavity, oral				3
xxxxxxx	2	R	M	1	498	1	1			BODT	body cavity, thoracic				3
xxxxxxx	2	R	M	1	498	1	1	ADEN	adenoma	GLPI	gland, pituitary	498	2	1	1
xxxxxxx	2	R	M	1	498	1	1			HEMO	hemolymphoreticular tissue				1
xxxxxxx	3	R	M	1	644	1	1			BODA	body cavity, abdominal				3
xxxxxxx	3	R	M	1	644	1	1			BODN	body cavity, nasal				3
xxxxxxx	3	R	M	1	644	1	1			BODO	body cavity, oral				3
xxxxxxx	3	R	M	1	644	1	1			BODT	body cavity, thoracic				3
xxxxxxx	3	R	M	1	644	1	1	ADEN	adenoma	GLPI	gland, pituitary	644	2	1	1
xxxxxxx	3	R	M	1	644	1	1			HEMO	hemolymphoreticular tissue				1
xxxxxxx	4	R	M	1	581	1	1			BODA	body cavity, abdominal				3
xxxxxxx	4	R	M	1	581	1	1			BODN	body cavity, nasal				3
xxxxxxx	4	R	M	1	581	1	1			BODO	body cavity, oral				3
xxxxxxx	4	R	M	1	581	1	1			BODT	body cavity, thoracic				3
xxxxxxx	4	R	M	1	581	1	1	ADEN	adenoma	GLPI	gland, pituitary	581	2	1	1

Appendix C: Business Rules for Datasets of Carcinogenicity Study

FDAB072	The microscopic examination findings for each organ/tissue should be consistent for each animal between the MI.xpt (SEND) and Tumor.xpt. Explain any terminology inconsistencies in the nSDRG.
FDAB073	The number of animals with a specific tumor finding for each organ/tissue should be consistent between the MI.xpt (SEND) and Tumor.xpt for the same dose group and sex. Explain any inconsistencies in the nSDRG.
FDAB074	The microscopic examination status of each organ/tissue should be consistent between the MI.xpt (MISTAT, MISPCUFL) and Tumor.xpt (ORGANEXM) for each animal. Explain any inconsistencies in the nSDRG.
FDAB075	For a specific organ/tissue, the number of animals that were examined (ORGANEXM=1), that were examined but not usable (e.g., autolyzed tissue or tissue not found in section examined) (ORGANEXM=2), and that were not examined (ORGANEXM=3) should be consistent between the MI.xpt (MISTAT, MISPCUFL) and Tumor.xpt for the same dosing group and sex. Explain any inconsistencies in the nSDRG.
FDAB076	The time in days/weeks to death or sacrifice and the status of death or sacrifice for each animal should be consistent between the SEND data and Tumor.xpt (DTHSACTM, DTHSACST). Explain any inconsistencies in the nSDRG.
FDAB077	For each specific tumor, the relationship to death in MI.xpt (MIDTHREL) should match with cause of death for the same tumor in Tumor.xpt (DEATHCAU). Explain any inconsistencies in the nSDRG.
FDAB078	For each specific tumor, the malignancy status should be consistent between the MI.xpt (MIRESCAT) and Tumor.xpt (MALIGNST). Explain any inconsistencies in the nSDRG.
FDAB079	The interval of time to detection of a tumor for each organ/tissue should be consistent between SEND (TFDETECT) and Tumor.xpt (DETECTTM). Explain any inconsistencies in the nSDRG.
FDAB080	The total number of animals assigned to each dose and sex group should be consistent between SEND and Tumor.xpt. Explain any inconsistencies in the nSDRG.
FDAB081	The tumor detection time of a specific organ/tissue/tumor should not be later than the death or sacrifice time for each animal.
FDAB082	For carcinogenicity studies, MIRESCAT should be populated unless MISTRESC has a value of 'UNREMARKABLE' (for SENDIG v3.1) or 'NORMAL' (for SENDIG v3.0) or MISTAT has a value of 'NOT DONE'.
FDAB083	If a specific organ was not examined in MI.xpt (MISTAT = 'NOT DONE'), then a record should be included in tumor.xpt for the same specific organ indicating that this organ was not examined (ORGANEXM=3).
FDAB084	If a specific organ was not examined in tumor.xpt (ORGANEXM=3), then a record should be included in MI.xpt (SEND) for the same specific organ indicating that this organ was not examined (MISTAT = 'NOT DONE').
FDAB085	For findings in MISTRESC using the NEOPLASM controlled terminology list, malignancy status in MIRESCAT should align with any benign or malignant designation in the NEOPLASM. Explain any inconsistencies in the nSDRG.

Appendix D: Mapping from SEND to tumor.xpt

Variable	Label	Type	Codes	SEND Notes [SEND Source Domain – Variable (CT*)]
STUDYNUM	Study number	Char		DM – STUDYID
ANIMLNUM	Animal number	Char		DM – USUBJID
SPECIES	Animal Species	Char	M = mouse, R = rat	DM – SPECIES
SEX	Sex	Char	M = male, F = female	DM – SEX
DOSEGP	Dose group	Num	Use 0, 1, 2, 3, 4,... In ascending order from control	TX – TXPARMCD TXVAL SETCD DM – ARM (TK status)
DTHSACTM	Time in days to death or sacrifice	Num		DS – DSSTDTC (Format - ISO 8601) EX – EXSTDTC (Format - ISO 8601) DSSTDTC - EXSTDTC + 1
DTHSACST	Death or sacrifice status	Num	1 = Natural death or moribund sacrifice 2 = Terminal sacrifice 3 = Planned intermittent sacrifice 4 = Accidental death	DS – DSDECOD
ANIMLEXM	Animal microscopic examination code	Num	0 = No tissues were examined 1 = At least one tissue was examined	MI – MISTAT (CT* - ND)
TUMORCOD	Tumor type code	Num		Not available in SEND
TUMORNAM	Tumor name	Char		MI – MISTRESC (CT - NEOPLASM)
ORGANCOD	Organ/tissue code	Char		Not available in SEND
ORGANNAM	Organ/tissue name	Char		MI – MISPEC (CT – SPEC)
MALIGNST	Malignancy status	Num	1 = Malignant 2 = Benign 3 = Undetermined	MI – MIRESCAT (CT – MIRESCAT)
DEATHCAU	Cause of death	Num	1 = Tumor caused death 2 = Tumor did not cause death 3 = Undetermined	MI – MIDTHREL (CT - NY)
ORGANEXM	Organ/tissue microscopic examination code	Num	1 = Organ/Tissue was examined and was usable 2 = Organ/tissue was examined but was not usable (e.g., autolyzed tissue) 3 = Organ/tissue was not examined	MI – MISTAT (CT – ND) and MISPCUFL (CT – NY) ORGANEXM=1 if MISTAT = null and MISPCUFL = null ORGANEXM=2 if MISPCUFL = "N" ORGANEXM=3 if MISTAT = "NOT DONE" and MISPCUFL = null
DETECTTM	Time in days/weeks of detection of tumor	Num	The number of days relative to the first day of treatment when the tumor was first detected	TF - TFDETECT or the following algorithm, if the tumor was ✓ detected as a clinical sign (CL): (CLDTC – EXSTDTC) + 1 ✓ detected as a palpable mass (PM): (PMDTC – EXSTDTC) + 1 ✓ detected during necropsy (MA): (DSSTDTC – EXSTDTC) + 1 ✓ detected during histopathology examination (MI): (DSSTDTC – EXSTDTC) + 1

* CT: Controlled Terminology