

Non-Clinical (SEND) Reference Guide for Clinical (SDTM) Programmers

Dharmendra Tirumalasetti, Vita Data Sciences, Waltham, USA;

Bhavin Busa, Vita Data Sciences, Waltham, USA

ABSTRACT

The U.S. FDA now requires the use of standardized data submission, SEND (Standard for Exchange of Nonclinical Data), for non-clinical data. Many Sponsor companies have started preparing SEND datasets towards their upcoming submissions, although, they still lack the much-needed expertise to get their data submission-ready. In our experience, one of the reasons could be due to lack of available resources/subject matter experts in the Non-clinical team within an organization. One of the solutions to overcome the resourcing challenges is to utilize existing pool of Clinical (SDTM) Programmers. In this paper, our intent is to provide a quick reference guide for Clinical (SDTM) Programmers to develop SEND domains for non-clinical studies. We will present commonalities and differences between SDTM and SEND domains. In addition, we will summarize our experience and lessons learned with performing mapping and standardization of non-clinical legacy studies to make it submission-ready.

INTRODUCTION

SEND has been required since at least 2017-12-17 for INDs and 2016-12-17. So, as to the question of “When do I first need to be SEND-ready”, the answer is “NOW”. The current version of SEND Implementation Guide is 3.1 which is based upon and should be used in close consent with Version 1.5 of the CDISC Study Data Tabulation Model (SDTM). The SDTM describes the conceptual model for representing study data for electronic data interchange and should be read prior to reading the SENDIG. The SENDIG provides specific domain models, assumptions, business rules, and examples for preparing standard nonclinical tabulation datasets that are based on the SDTM. This version of the SENDIG is designed to support data typically found in single-dose general toxicology, repeat-dose general toxicology, and carcinogenicity studies, as well as respiratory and cardiovascular testing done during safety pharmacology studies. However, the SENDIG can also be used to represent data for other study types.

DETAILS

COMMONALITIES AND DIFFERENCES BETWEEN SEND AND SDTM

The SDTM is built around the concept of observations collected about subjects included in a study. The Implementation Guide for SEND is based on the SDTM general framework for organizing nonclinical study information that is to be transferred between parties or submitted to regulatory authorities. Test results, examinations, and observations for subjects in a nonclinical study are represented in a series of SEND domains. Although SENDIG V3.1 has been developed based on the SDTM version 1.5, there are many differences between SENDIG V3.1 and SDTMIG V3.2 that we should keep in mind while developing SEND datasets. The following figure provides the list of common domains between SENDIG V3.1 and SDTMIG V3.2 along with the specific domains to SEND and SDTM only.

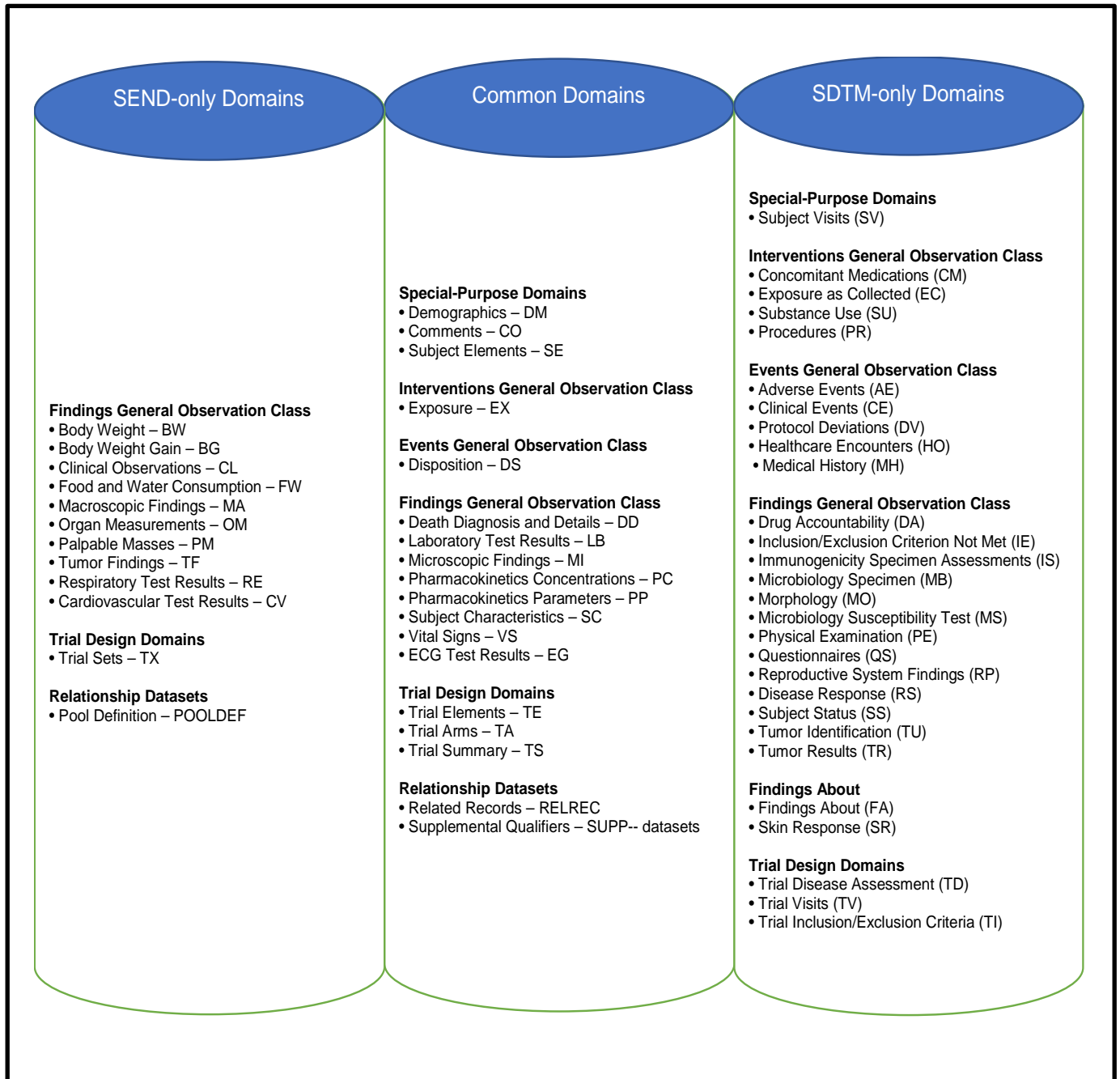


Figure 1. Commonalities and differences between SDTMIG V3.2 and SENDIG V3.1 Domains.

The programmers who work on developing the SEND domains should keep in mind that the domains which are specific to SDTM should not be created as part of SEND. In addition to the above list of differences, there is a list of variables which should not be used in SEND when compared to SDTM.

Demographics	Events	Interventions	Trial Summary
<ul style="list-style-type: none"> • ACTARMCD • ACTARM • COUNTRY • DTHDTC 	<ul style="list-style-type: none"> • BDSYCD • HLT • HLTCD • HLGT • LLT 	<ul style="list-style-type: none"> • PRESP 	<ul style="list-style-type: none"> • TSVALCD • TSVCDREF • TSVCDVER

Demographics	Events	Interventions	Trial Summary
<ul style="list-style-type: none"> • DTHFL • ETHNIC • INVID • INVNAM • RACE • RFICDTC • RFPENDTC 	<ul style="list-style-type: none"> • LLTCD • PARTY • PRYID • PTC • SCAN • SCONG • SDISAB • SDTH • SHOSP • SLIFE • SOD • SMIE • SOC • SOCCD 		

Table 1. List of variables that should not be used in SEND Datasets.

Similarly, there are some variables which are specifically used in SEND datasets only such as:

- --UNSCHFL Unscheduled Flag, in Events, Interventions and Findings Domains.
- --IMPLBL Implantation Site, in Findings Domains.
- --FETUSID Fetus Identifier, in all classes as identifier variable.
- --NOMDY Nominal Study Day, Timing variable in the applicable domains.
- --NOMLBL Label for Nominal Study Day.

In non-clinical trials, a single finding can be determined or calculated at a level of pool of subjects instead of on one subject. For example, some clinical findings are determined for a group of animals put in a cage instead of collecting the results for each animal. In these scenarios, POOLID should be populated and USUBJID in any of the findings domains is blank. Associated POOLDEF dataset needs to be created with variables POOLID and USUBJID populated in a study.

As development of SDTM domains has been widely used already for the data collected on CRFs in an EDC format, the data acquisition for clinical trials has many standard approaches and databases have been developed for the clinical data to be stored. Many sponsors have also started implementing CDASH standards for acquisition of clinical data.

In our experience of working on non-clinical data, as implementation of SEND is fairly new, the source data on animals has been captured in file formats of excel, csv, text etc. Many sponsors or vendors captured the non-clinical data with-out following any guidelines; especially comments & clinical observations have been captured manually. Sometimes data has been collected for each animal which has clinically significant results separately where it is a single comment for all the animals together in some cases. This has become a challenge and we had to request our clients to make sure to collect the data at animal level so that this can be used in the development of SEND domains. We also proposed the templates to convert the data captured into animal/pool level data that is required for the submission of non-clinical data in SEND format.

The following is the expected list of SEND domains, along with define.xml, that should be generated as part of submission for a non-clinical trial:

- DM (Demographics)
- DS (Disposition)
- EX (Exposure)
- TA (Trail Arms)
- TE (Trail Elements)
- TX (Trial Sets)
- TS (Trial Summary)

In addition to the SEND domains, it is recommended to read and know about **tumor.xpt** file.

The tumor.xpt file is a nonclinical analysis dataset. Specific assumptions regarding tabulating data from carcinogenicity studies were created with the intent that SEND datasets will contain the necessary information to derive a tumor.xpt file. It is the intent of the CDISC SEND team that the data meet regulatory needs if submitted as SEND datasets.

The following assumptions must be met in to create a tumor.xpt file:

1. Every subject must have at least one record in the EX (Exposure) domain with EXSTDTC populated.
2. All organs scheduled for examination must have a record in the MI (Microscopic Findings) domain even if they were not analyzed. If a scheduled tissue is not examined, then a record for that tissue should be included with MISTAT= "NOT DONE". In any instance where a sample is found unusable (i.e., autolyzed), MISPCUFL must be "N."
3. If a sponsor chooses to include secondary or multicentric tumors in the tumor.xpt, they must map METASTATIC to MALIGNANT (=1) for the MALIGNST variable.
4. If a sponsor does not choose to include secondary or multicentric tumors in the tumor.xpt, the sponsor must exclude records with TFRESCAT= "METASTATIC."
5. Every tumor (including secondary and/or multicentric tumors) must have one record in the TF (Tumor Findings) domain.
6. Secondary tumors must have a MIRESCAT value of "METASTATIC."
7. When creating the Tumor Findings domain, all secondary tumors must contain a value of "MALIGNANT" in the TFRESCAT variable.
8. TFDETECT (Time in Days to Detection of Tumor) is the number of days from the start of dosing to the earliest detection of the tumor in the experimental.

The following domains are required in order to create a tumor.xpt file:

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

STUDY DATA STANDARDIZATION PLAN

For clinical and nonclinical studies, sponsors should include a plan (e.g., in the IND) describing the submission of standardized study data to FDA at the IND time point. The Study Data Standardization Plan (SDSP) also establishes and documents a plan for describing the data standardization approach for nonclinical studies (SEND) and clinical (SDTM/ADaM) within a development program. The Study Data Standardization Plan (SDSP) assists FDA in identifying potential data standardization issues early in the development program. Sponsors may also initiate discussions at the pre-IND stage. For INDs, the SDSP should be located in the general investigational plan.

The development and maintenance of the Study Data Standardization Plan provides important benefits to sponsors. First, the SDSP brings internal focus and agreement to the standards throughout the project/program lifecycle versus an ad hoc process driven by individual studies. Second, the availability of the SDSP provides an important reference for all sponsor groups (e.g., regulatory affairs, data management). Further this document provides a means of tracking discussions and agreements with the FDA. Finally, this document will drive decisions about legacy data conversion and up-versioning of data standards to allow for pooling of data across studies at a much earlier time than pre-NDA meetings.

The SDSP should be updated in subsequent communications with FDA as the development program expands and additional studies are planned. Updates to the SDSP should not be communicated each time a study is started. The cover letter accompanying a study data submission should describe the extent to which the latest version of the SDSP was executed. In addition, for clinical studies that will be submitted to

CBER, the SDSP appendix should be provided to the review office no later than the end-of-phase 2 meeting. The CBER SDSP appendix should include tables of proposed SDTM domain/variable usage, supplemental domain usage and proposed analysis.

LESSONS LEARNED

As mentioned earlier that SEND is fairly new and not all the sponsors have standardized way of collecting or acquiring data for non-clinical studies, it becomes challenging to convert the raw/source data into SEND as no guidelines have been followed. It is always recommended to have cross-functional discussions to make the DM team of non-clinical data understand the requirements at the beginning of non-clinical study itself. It is recommended to involve the programming team along with submissions team in the discussions. This way, it would become easier for the programmers to convert these source data, if collected at standard format, into SEND format. Any pool level information is collected, it is important to know and collect the subject IDs which are part of each pool/cage.

CONCLUSION

As many sponsor companies started developing SEND domains for their upcoming submissions, it is recommended to involve the programming lead or manager at the start of design of the data collection forms and creation of database for non-clinical studies. This would help the teams to understand how the non-clinical data need to be collected or acquired and programming teams would be able to convert the source data on animals to SEND domains without much hassles.

All the SAS® programmers need to keep in mind that there are many differences between SEND and SDTM even though they both follow the same SDTM model. Before working on any SEND domain, programmer needs to know the variables that can not be part of SEND domains. Provided the clinical programmers know the nuances of SEND and differences or commonalities between SEND and SDTM, Sponsor successfully can utilize the clinical programmers to work on and generate SEND domains for submission.

REFERENCES

[1] "CIDSC SDTM Model V1.5" Available at <https://www.cdisc.org/standards/foundational/sdtm/sdtm-v15>

[2] "SENDIG V3.1" Available at

<https://www.cdisc.org/standards/foundational/standard-exchange-nonclinical-data-send/sendig-v31>

[3] "SDTMIG V3.2" Available at <https://www.cdisc.org/standards/foundational/sdtmig/sdtmig-v32>

[4] "STUDY DATA TECHNICAL CONFORMANCE GUIDE – March 2018." Available at

<https://www.fda.gov/media/88173/download>

[5] "Study Data Standardization Plan (SDSP) - Sponsor Implementation Guide" Available at

<https://www.phuse.eu/documents//sop/wp/phuse-wp002-sdsp-sponsor-implementation-guide-v1-8408.pdf>

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Name: Dharmendra Tirumalasetti

Company: Vita Data Sciences

Address: 281 Winter Street, Suite 100

City / Postcode: Waltham, MA 02451

Phone: 781-833-0257

Email: dtirumalasetti@vitadatasciences.com

Web: www.vitadatasciences.com

Name: Bhavin Busa

Company: Vita Data Sciences

Address: 281 Winter Street, Suite 100

City / Postcode: Waltham, MA 02451

Phone: 781-373-8455

Email: bbusa@vitadatasciences.com

Web: www.vitadatasciences.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration. Other brand and product names are trademarks of their respective companies.