

Qualitative Analysis of Actual SEND Datasets for Trial Design Domains

G. Sato^{1,2}, M. Suzuki^{1,3}, M. Nakajima^{1,4}, T. Kodama^{1,5}, N. Ohara^{1,6}, T. Tanaharu^{1,7}, A. Kiya^{1,8}, T. Kinomoto^{1,9}, M. Yoshiike^{1,10},

I. Matsumoto^{1,11}, T. Isobe^{1,12}, K. Hara^{1,13}, T. Yamaguchi^{1,14}, Y. Nakano^{1,15}, K. Motoyama^{1,16}, T. Hamano^{1,17}, K. Watanabe^{1,12}

¹SEND Taskforce (TF) Team, Non-clinical Evaluation Expert Committee, Japan Pharmaceuticals Manufacturers Association, ²Eisai Co., Ltd., ³Kyowa Hakko Kirin Co., Ltd., ⁴Asahi Kasei Pharma Corporation, ⁵EA Pharma Co., Ltd., ⁶Otsuka Pharmaceutical Co., Ltd., ⁷Sanwa Kagaku Kenkyusho Co., Ltd., ⁸Shionogi & Co., Ltd., ⁹Zeria Pharmaceutical Co., Ltd., ¹⁰Daiichi Sankyo Co., Ltd., ¹¹Sumitomo Dainippon Pharma Co., Ltd., ¹²Chugai Pharmaceutical Co., Ltd., ¹³Nippon Kayaku Co., Ltd., ¹⁴Eli Lilly Japan K.K., ¹⁵Mochida Pharmaceutical Co., Ltd., ¹⁶Janssen Pharmaceutical K.K., ¹⁷Novartis Pharma K.K.

Disclaimer: The opinions and recommendations in this presentation results from discussions of the SEND Taskforce team, Non-clinical Evaluation Expert Committee, JPMA, and cannot be used without prior notice to the author(s) and appropriate citation. Correspondence: g-sato@hcc.eisai.co.jp

Introduction

In December 2017, preparation and submission of SEND datasets became a regulatory mandate for certain types of nonclinical toxicology studies (i.e. single and repeat-dose general toxicology, and carcinogenicity studies) for IND applications in addition to NDA submissions in the United States. As of February 2018, the PMDA has not decided whether they will require SEND datasets for regulatory submissions in Japan, thus many Japanese pharmaceutical companies, especially those who mainly target the domestic market, have not fully implemented a process for SEND submission to regulatory authorities. The current situation in Japan is that, since April 2017, whether to introduce SEND in Japan has been a topic of discussion at AMED research groups among industry, government, and academia in Japan; and global companies who target US market took the initiative in implementing SEND.

To further understand implementation of SEND, SEND Taskforce (TF) Team, Non-clinical Evaluation Expert Committee, JPMA started a project to learn what kind of information should be populated in which domain in what way. It was anticipated that achieving our goals by simply reading SENDIG v3.0³ was not easy, thus we decided to examine the information in the contents of actual SEND datasets in order to learn details such as how to populate variables and parameter codes.

Methods

The SEND TF Team conducted a series of surveys to collect information on the contents of actual examples of SEND datasets owned by the Team members (i.e. 16 pharmaceutical companies) from August 2017 to February 2018. The TF Team focused on Special-Purpose (DM and SE) and Trial Design (TE, TA, TX, and TS) domains that were compliant with CDISC SENDIG v3.0. The reason for focusing on these domains was that a toxicologist who actually runs a toxicology study may have more concerns describing Study Schedule in SEND datasets compared with dealing with study results. The survey included 2 questionnaires: one to understand whether individual variables were populated in each domain or parameter code in TX/TRIAL SET and TS/TRIAL SUMMARY, and the other to understand how the variable was populated. Team members were required to replace some of the contents of their SEND datasets with other valid characters for confidentiality purposes and to maintain blinded qualitative analysis, prior to submission to the TF.

Results

The number of answers to the questionnaires collected from Team members, or in other words, the number of studies reported was 13 to 14 (with one exception) from among the 16 companies. Note that we separately conducted surveys on individual domains, therefore the number of responses differed among domains, thus the component of the responses may be different among domains. Some companies could not provide the answers since they had no experience in creating SEND datasets, while others provided information from more than one study. The TF Team did not confirm if the SEND datasets included in the survey had been submitted to and/or accepted by the FDA.

The section below describes, 1) whether an individual variable/parameter code was populated or not, and 2) how the variable/parameter code was populated as well as "Points to Consider" when populating variables/parameter codes. Additionally, there was a high degree of variability in values describing study schedule, test articles and test methods in SEND, which were also described in 2).

1) Whether to Populate Each Variable/Parameter Code

Whether to populate each variable/parameter code was shown for each domain in Fig. 1. The percentages for populating variables whose Core=Req (Required) or Exp (Expected), or parameter codes that are "Should Include – Yes" were almost 100% in TRIAL SET and TRIAL SUMMARY, while those for populating variables whose Core=Perm (Permissible) or parameter codes that are NOT "Should Include – Yes" had large variation from 0% to 100%. Especially the parameter codes in TX/TRIAL SET and TS/TRIAL SUMMARY were not often populated (<50%) in general if they were not "Should Include – Yes". For example, animal husbandry-related parameter codes (e.g. temperature [ENVTEMP], humidity [HUMIDT], light [LIGHT], housing [HOUSEHYP], bedding [BEDDING], diet [DIET], water [WATER]) were rarely populated (<30%) in TS/TRIAL SUMMARY. It should be noted that the fact is, collected information may not always be submitted via SEND, although the Section 4.1.4.4.4 of the TCG v4.0³ says, "For the purposes of SDTM and SEND submissions, all Required, Expected, and Permissible variables that were collected, plus any variables that are used to compute derivations, should be submitted".

2) How to Populate the Variable/Parameter Code – Points to Consider

We found variations in populating some variables/parameter codes in this survey. The relationship between items described in Final Study Reports and variables that were populated in SEND was summarized in Table 1, and the variations identified in this survey were described as "Points to Consider". Important points identified during team discussions are also included under "Points to Consider", even if there was no variability.

There were marked variations in populating some Study Design-related variables, such as EPOCH, ETCDELEMENT, ARMCD/ARM, SETCD/SET, which were roughly categorized into a few groups. Additionally, considerable variations were found in values included in Timing variables (e.g. SESTDTC/SEENDTC in SE domain; see below). Variations in values for describing test articles and test methods were also found.

With regard to Timing variables, there is a variation in describing Duration, i.e. whether there are time gaps between Elements. In 9 out of 14 examples (64%), the value of SEENDTC for one ELEMENT is exactly same as the value of SESTDTC for the next ELEMENT. The remaining 5 examples only contain year, month, and date (YYYY-MM-DD), and 1-day difference was seen between the SEENDTC and the next SESTDTC values, which was considered out of the rules/assumptions in SENDIG v3.0. Note that SENDIG v3.1 allows the value of SEENDTC for one Element immediately before the value of SESTDTC for the next Element, while SENDIG v3.0 does not allow time gaps between Elements. Therefore, it is suspected that some of the SE domains collected in this survey actually followed SENDIG v3.1 in advance of the requirement.

Discussion

There was a presentation about whether an individual variable was populated in each domain or parameter code in TX/TRIAL SET and TS/TRIAL SUMMARY 3), however, as far as we know, this was the first presentation that considers how to populate each variable based on the information of actual SEND datasets. First, the SEND TF Team would like to emphasize that the primary purpose of this survey is to reveal the current situation, and NOT to decide which variables do not need to be populated by a majority vote.

As expected, this survey revealed variations in some variables and parameter codes among SEND datasets. There was a high degree of variability in populating variables/parameter codes for which a CT codelist does not exist. Such variations were mostly found in variables such as ARM/ARMCD, SET/SETCD, ELEMENT/ETCD, EPOCH, which are "SEND-specific" concepts ("CDISC standard-specific" may be more appropriate) and which are not familiar to many preclinical toxicologists. For example, a wide variety of values (Screen, Screening, Acclimation, Predosing, Pre-dosing, Predose, Pretreatment) was included for Pretreatment phase of EPOCH in TA domain. Variation in naming rules for such variables became apparent because a CT codelist does not exist for them and also as they should be populated as most of their Core = Req or Exp.

When outsourcing SEND dataset creation it is important to pay attention to variations in populating variables/parameter codes, as mentioned above, which may differ among CROs/authors. As the SENDIG v3.0, Section 5.1.1.1 states: "The assignment of the values of ARMCD and ARM should be consistent, if possible, within a submission". It is the Sponsor's responsibility to pre-determine the naming rules for such variables and insure consistency across SEND datasets within a submission package for a new drug. Otherwise the above-mentioned variations could be an issue if the package contains SEND datasets that are created with contributions by different authors.

Next, we considered that variation in SEND datasets may reflect the differences in Final Study Report (e.g. content and format following SOPs). We listed the items that should be included in the Final Study Reports as defined in GLP guidelines of the FDA, OECD, and Japan/MHLW and tied with some typical variables in SEND datasets in Table 1. For example, even if the test compound is the same, a Sponsor may face variations in the name of the vehicle for the compound that could, for example, be described as either %GLUCOSE or % Glucose in a Final Study Report depending on the test facility or CRO; the TRTV value in the SEND datasets usually reflects the terminology used in the Final Study Report and therefore vary depending on the description. The name of the diet; DIET value in the SEND datasets can be either CRF-1 (actual brand name) or STANDARD, which is another example of variation. These cases are just examples, and some other variables that do not have CT may be populated in the same way. If general naming rules within a Sponsor could be established, although naming of certain variables may be limited depending on the LIMS system used at the facility/CRO, this content would become more useful in the searchability in accumulated data within the Sponsor.

Conclusion

This survey conducted by the JPMA SEND TF Team, although a relatively small scale survey, successfully revealed several issues in SEND datasets created according to the SENDIG v3.0. The qualitative analysis of the results indicated that variations exist based on whether or not an individual variable/parameter code was populated, and on how the variable/parameter code was populated. Although it could take some time, the TF Team members consider that reducing such variations should lead to higher quality datasets with powerful and increased searchability so that accumulated SEND datasets should become more valuable, and recommend this approach. These efforts would provide regulatory agencies with easier review of SEND datasets, which can accelerate new drug approval. However, the TF Team also recommends reducing variations in the description of the Final Study Reports, otherwise the value of consistency efforts will be reduced. Furthermore, the TF Team emphasizes the current urgency of the opportunity for creating SEND datasets, and for Sponsors, IT vendors, CROs, and Regulatory Agencies to work together to reduce variations in the Final Study Reports as well as in the SEND datasets. Since SEND dataset creation and submission is mandated for US IND application submission, these recommendations will insure that the SEND dataset will truly contribute to drug safety assessment.

Acknowledgements

The authors would like to thank all the reviewers in the TF Team members' companies for their thorough review and useful suggestions in poster preparation. They also gratefully acknowledge Ms. Megan E. Bausman and her colleagues for their valuable comments to improve the quality of the poster as experts in one of the most experienced CROs in SEND implementation in the world.

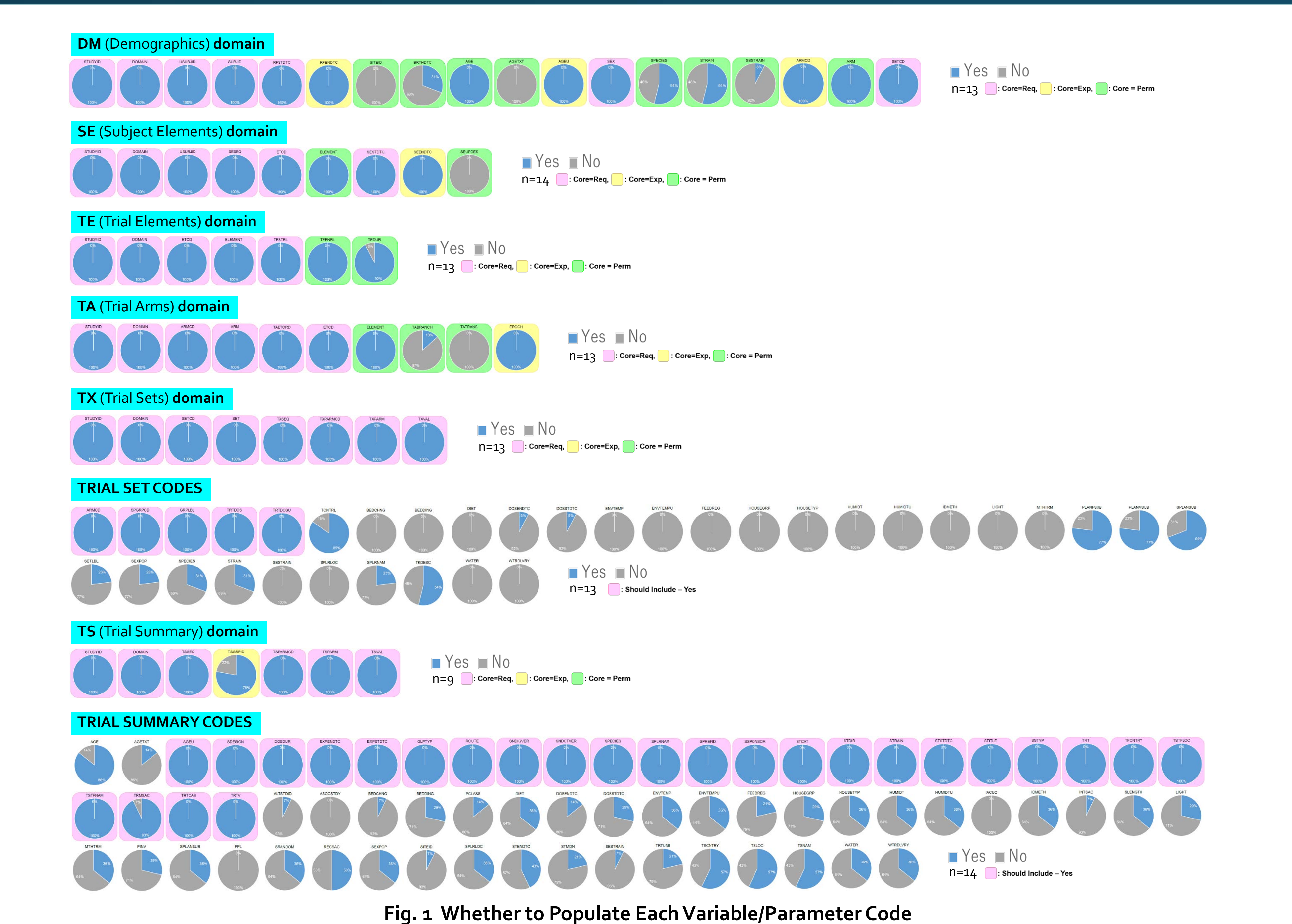


Fig. 1 Whether to Populate Each Variable/Parameter Code

References

1) Clinical Data Interchange Standards Consortium (CDISC) (2011) Standard for Exchange of Nonclinical Data Implementation Guide: Nonclinical Studies Version 3.0. 2) U.S. Food and Drug Administration, Center for Drug Evaluation and Research (CDER)/Center for Biologics Evaluation and Research (CBER) (2017) Study Data Technical Conformance Guide Version 4.0. 3) G. Sato, Y. Sakakibara, A. Hamada, T. Uota, M. Yoshiike, Y. Fujimura. Differences in SEND Dataset Creation: Analysis of Current Situation and Points to Consider. The Toxicologist Volume 166, Issue 1, March 2017, S279. 4) Electronic Code of Federal Regulations, Title 21 Food and Drugs (2018) https://www.ecfr.gov/cgi-bin/text-idx?SID=3ee286322a6f6a39d9e6d786a60a8b&mc=true&tpl=/ecfrbrowse/Title21/21tab_02.tpl Accessed February 20 2018. 5) OECD SERIES ON PRINCIPLES OF GOOD LABORATORY PRACTICE AND COMPLIANCE MONITORING Number 1. OECD Principles on Good Laboratory Practice (as revised in 1997). ENV/MC/CHEM(98)17. (1998) [http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?docLanguage=en&code=en/mc/chem\(98\)17](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?docLanguage=en&code=en/mc/chem(98)17) Accessed February 20 2018. 6) Ministerial Ordinance on Good Laboratory Practice for Nonclinical Safety Studies of Drugs. Ordinance of the Ministry of Health and Welfare No.21 of March 26, 1997 (as last amended by the Ordinance of the Ministry of Health, Labour and Welfare No.114 of June 13, 2008). (2008) <https://www.pmda.go.jp/files/000204103.pdf> (in Japanese). Accessed February 20 2018.

Abbreviations

AMED = Japan Agency for Medical Research and Development, CRO = Contract Research Organization, CT = Controlled Terminology, DM = Demographics, IND = Investigational New Drug, JPMA = Japan Pharmaceuticals Manufacturers Association (Japanese PhRMA), LIMS = Laboratory Information Management System, MHLW = Ministry of Health, Labour and Welfare, MHW = Ministry of Health and Welfare, NDA = New Drug Application, PMDA = Pharmaceuticals and Medical Devices Agency, SEND = Standard for Exchange of Nonclinical Data, SE = Subject Elements, SENDIG = SEND Implementation Guide, TA = Trial Arms, TCG = Technical Conformance Guide, TE = Trial Elements, TS = Trial Summary, TX = Trial Sets.

Table 1 Relationship between Items in Final Study Reports and Some Typical Variables/Parameter Codes in SEND Datasets (CT version referred: 2017-09-29*)

GLP**	Items in Final Report	Corresponding Variables/Parameter Codes in SEND Datasets	Points to Consider
#	Study Title	STITLE**	No issues detected for Study Title.
#	Study Number	STUDYID** SPREF**	No issues detected for Study Number.
#	Objectives	STCAT** SSTYP** SDSIGN** SRANDOM	Full sentence of Objectives does not need to be populated in the SEND Datasets; only information on Study Category, Study Type, Study Design, whether the "Study is Randomized" are required.
#	Name and Address of the Test Facility	Test Facility Name Test Facility Location Test Facility Country	No issues detected for Name and Address of the Test Facility.
#	Name and Address of the Test Site	Test Site/Vendor/Laboratory Name Test Site Location Test Site Country	No issues detected for Name and Address of the Test Site.
#	Name and Address of the Sponsor	Sponsoring Organization SSPONSOR**	No issues detected for Name and Address of the Sponsor.
#	Study Initiation/Completion Date	Study Initiation Date Study Completion Date	No issues detected for Study Initiation/Completion Date.
#	Experimental Starting/Completion Date	Experimental Starting Date Experimental Completion Date	EXPSTDTC, EXPENDTC. If OECD-GLP is not applied for a study, the definition of Experimental Period may differ among facilities.
#	Name of the Study Director	Study Director STDIR**	No issues detected for Name of Study Director.
#	Name of the Principal Investigator(s), and the Phase(s) of the Study Delegated	Principal Investigator PINV	Information does not need to be populated in the SEND Datasets.
#	Name(s) of the Other Scientists or Professionals	Evaluator Sponsor's Monitor Contributing Scientist	No issues detected for Name(s) of the Other Scientists or Professionals.
#	Test and Control/Reference Articles	Investigational Therapy or Treatment Name of Actual Treatment Primary treatment CAS Register Number Pharmacologic Class Primary Treatment Unique Ingredient ID Lot Number Control Type Treatment Vehicle Treatment Vehicle	TRT** EXTRT** TRTCAS** PCLASS TRFUNII EXLOT TXCTRL EXTRV PCLASS. We consider it is important to indicate the pharmacologic class to analyze class effect using accumulated data, for example. However, the actual Study Protocol/Final Study Report usually does not contain such a information. TRTV. Since there is no CT exist, each facility may decide the vehicle name at its sole discretion and which leads to variations among facilities (e.g. %GLUCOSE vs. % Glucose).
#	Stability (and Homogeneity) of the Test and Control/Reference Articles under the Conditions of Administration	Test System Used	Information does not need to be populated in SEND Datasets.
#	GLP and Other Test Guideline(s)	Good Laboratory Practice Type GLP Flag, Good Laboratory Practice Flag	GLPTYP** GLPLF** There may be some discussion how to describe Japanese GLP (i.e. MHW vs. MHLW; when the current J-GLP was enforced in 1997 the ministry's name was MHW, but it was changed to MHLW in 2001).
#	Statistical Methods	Deviations from the Study Plan	Information does not need to be populated in the SEND Datasets.
#	Results	Description of Unplanned Element Examination Status Reason Not Done	SEUPDES --STAT (CL, MA, MI etc.) --REASND (CL, MA, MI etc.) Deviations are reflected to the SEND Datasets if categorized as Findings General Observation Class, while they will not if related to animal husbandry (e.g. Environmental Temperature, Housing Humidity, etc.)
#	Location(s) Where Raw Data, All Specimens, and the Final Report are to be Stored	Route of Administration Dose per Administration Dose Level Dose Unit Dosing Duration Dose Form Dosing Frequency Per Interval Start Date/Time of Dose Interval End Date/Time of Dose Interval Study Length Time to Interim Sacrifice Recovery Period Test Name Test Short Name Method of Test or Examination Specimen Material Type (Specimen Type) Fasting Status	ROUTE** (EXROUTE**) EXDOSE TRTDOS** TRTDOSU** (EXDOSU) DOSDUR** EXDOSFRM EXDOSFRQ DOSSTDTC DOSENDTC SLENGTH INTSAC --TRTNAME** RECSAC --TEST** (LB, MI etc.) --TESTCD** (LB, MI etc.) --METHOD (LB, MI etc.) --SPEC** (LB, MI etc.) --FAST (BW, LB, PC) ROUTE, EXROUTE. Some of the CT (ROUTE) are considered similar and exchangeable (e.g. INTRAVENOUS and INTRAVENOUS BOLUS, ORAL and ORAL GAVAGE). TRTDOS. There are two patterns; one is simply dose level only (e.g. 0, 10, 30, 100) and the other is with group name (e.g. Group 1 - 0, Group 2 - 10, Group 3 - 30, Group 4 - 100). TRTDOSU. Selection of Dose Unit may vary among facilities, "mg/kg" and "mg/kg/day" are most commonly used in a toxicology study. DOSDUR. Dosing Duration of a 1-month toxicity study could be expressed as P28D, P4W, or P2M; it may depend on the facility's rule. TRTNAME. Although TRTNAME is a variable which is "Should Include – Yes" and populated in ISO 8603, it could be null or "Not Available" if the test subjects are returned to a colony alive.
#	Reports of Principal Investigators or Scientists	Study Director Statement	Information does not need to be populated in the SEND Datasets.
#	QA Statement	QA Statement	Information does not need to be populated in the SEND Datasets.
#	Amendments	Amendments	If the Study Design is amended, the new Study Design should be reflected to the SEND Dataset.
#	Study Date/Day, Animal Number, Study Design Related and Other (SEND Specific Concept)	Date/Time of -- Study Day of -- Planned Time Point Name Planned Elapsed Time from Time Point Ref Unique Subject Identifier Domain Abbreviation Description of Planned Arm Planned Arm Code Set Description Set Code Set Label Description of Element Element Code Order of Element within Arm Rule for Start of Element Rule for End of Element Start Date/Time of Element End Date/Time of Element Planned Duration of Element Trial Epoch Sponsor-Defined Group Code Group Label Toxicokinetic Description Sequence Number Group Identifier SEND Implementation Guide Version SEND Controlled Terminology Version	--DTC --DY --TPT --ELTM USUBJID** DOMAIN** ARM** ARMCD** SET** SETCD** SETLBL ELEMENT** ETCD TACTORD** TESTRL** TEENRL** SESTDTC** SEENDTC** TEDUR EPOCH SPGRP** GRPLBL** TKDESK --SEQ** (MA, MI etc.) --GRPID (MA, MI, TS etc.) SNOIGVER** SNOICTVER** USUBJID. There are at least two different naming rules (i.e. STUDY#-ANIMAL# and STUDY#-ANIMAL#). ARM, ARMCD. There are wide variations in naming rule; it is important to understand that the assignment of values to ARMCD/ARM should be consistent, if possible, within a submission (SENDIG v3.0, Sec. 5.1.1.1). SET, SETCD. There are wide variations in naming rule as seen in the ARM/ARMCD. ELEMENT, ETCD. There are wide variations in naming rule, especially for ETCD. TESTRL, TEENRL. There are some variations in the description of TESTRL and TEENRL. SESTDTC, SEENDTC. According to the SENDIG v3.0, there are NO TIME GAPS between Elements; however, SENDIG v3.1 allows to have the value of SEENDTC immediately before the value of SESTDTC for the next ELEMENT. This type of variation is actually found in our analysis, even though the collected Datasets were created according to the SENDIG v3.0. EPOCH. There are wide variations in naming rule especially for Pretreatment phase (e.g. Screen, Screening, Acclimation, Predosing, Pre-dosing, Pre-Dosing, Predose, Pretreatment). SETLBL, SPGRP, CD, GRPLBL. There are some variations in naming rule in these variables.

XXXXXX. See "Points to Consider"

**If description in SENDIG v3.0 and CT 2017-09-29 is different, that in CT is prioritized (e.g. CLASS vs. PCLASS).

**Indicate if the item should be described in the Final Report under the GLP (F: FDA-GLP, O: OECD-GLP, J: Japan/MHLW-GLP).

**Although this information may be described in the Final Report, there is no corresponding item in the corresponding GLPs.

**Core = Req or Should Include – Yes (in either of the domain(s)), for variable which presents in more than one domain).

***Not listed in the SENDIG v3.0 but in the CT (2017-09-29).