

Stay ahead of the game: Know the SDTM compliance check reject messages before creating your SDTM datasets

Lixiang Larry Liu, Yuqin Helen Li, Eli Lilly and Company, Indianapolis, IN

ABSTRACT

Beginning October 2016, Pharmaceuticals and Medical Devices Agency (PMDA) has started accepting electronic study data submissions with a 3.5 year transitional period. Even though PMDA's requirements are very similar to those of U.S. Food and Drug Administration (FDA), there are some differences between these two regulatory agencies. One of the key differences is the severity level for both the SDTM and ADaM validation rules. While FDA only has the error and warning messages, PMDA also has the reject severity level, and has clearly stated that "Rules which, if violated, will cause the review to be suspended until corrections are made." Because the drug development is done globally, and generally, the same submission package will go to both FDA and PMDA, the study team should avoid all the reject messages in both SDTM and ADaM submission packages.

This paper provides a comprehensive review of the nine reject SDTM validation rules per PMDA severity level. Having a clear understanding of these rules can help a study team to stay ahead of the game and to avoid non-compliance starting from the SDTM specification, which will enable the team to build a solid foundation for the submission.

INTRODUCTION

Pharmaceuticals and Medical Devices Agency (PMDA) has started accepting electronic study data submissions with a 3.5 year transitional period since October 2016. One of the key differences between FDA and PMDA's validation rules is the severity level for both the SDTM and ADaM. While FDA only has the error and warning messages, PMDA also has the reject severity level, and has clearly stated that reject messages are not allowed in the submission package. Given the same submission package will go to both FDA and PMDA for most clinical trials, study teams should avoid all the reject messages in both the SDTM and ADaM submission packages.

Statistical group should take the initiative and stay ahead of the game by knowing what the reject validation rules are at the beginning of their study, so that they can try to avoid of them starting from the SDTM specification (This paper will only focus on the SDTM compliance check due to the space). The last thing the study team wants to see is a reject message from the compliance check after the database lock, making the team scramble to dig out the root cause of the issue and spend time and effort to fix it. In the worst case scenario, a re-lock might be necessary in order to remove the reject message.

In the following sections, this paper provides a comprehensive review and in-depth look of all the nine reject validation rules per PMDA standard, which will enable the readers to have a better understanding of these rules, and avoid of these reject messages in their submission packages.

Reject message #1: Variable value not found in non-extensible codelist

This is the validation rule of FDAC340, PMDA CT2001, and it applies to all SDTM domains. The description for this validation rule is :“Variable must be populated with terms from its CDISC controlled terminology codelist. New terms cannot be added into non-extensible codelists.”

Some controlled terminologies (CT) have non-extensible codelists, while other CTs have extensible codelists. This validation rule is saying that for some of the non-extensible codelists CT, you can not add a new value to its current list of values. For example, variable AESER in AE domain, which is the variable to indicate whether it is a serious adverse event or not, the submission value for AESER could only be Y or N, nothing else. If you assign a new value to this variable, you will get the afore mentioned reject message in your compliance check report.

PMDA further noted “Violation of CT2001 for the following variable/codelist pairs will be considered part of Rejection criteria. All others will be considered Errors.”

Variable Name (Codelist Name)

AGEU (AGEU)

COUNTRY (COUNTRY)

IECAT (IECAT)

RELTYPE (RELTYPE)

SEX (SEX)

--BLFL (NY)

--DRVFL (NY)

--FAST (NY)

--OCCUR (NY)

--PRESP (NY)

AESER (NY)

AESCONG (NY)

AESDISAB (NY)

AESDTH (NY)

AESHOSP (NY)

AESLIFE (NY)

AECONTRT (NY)

AESCAN (NY)

AESMIE (NY)

AESOD (NY)

IEORRES (NY)

IESTRESC (NY)

--STAT (ND)

Please note for the above 23 variables, some of the variable names might vary depending on the SDTM domain. For example, for the --STAT (ND) variable, it is LBSTAT in LB domain, and QSSTAT in QS domain. PMDA is saying explicitly that for all the above 23 variables, you can't add a new value to its current list of values, otherwise, you will get a reject message. For all other non-extensible codelists CTs, e.g. AEOU, AEACN, etc., if you add a new value, you would get an error message, which is less severe than reject message, but it is still not desirable.

Reject message #2: Variable value not found in non-extensible codelist when value-level condition occurs

This is the validation rule of FDAC343, PMDA CT2004, and it only applies to Questionnaire (QS) and Trial Summary (TS) SDTM domains. The description for this validation rule is :“Variable must be populated with terms from its CDISC controlled terminology codelist, when its value level condition is met. New terms cannot be added into non-extensible codelists.”

PMDA further noted :“Violation of CT2004 for the following variable/codelist pairs will be considered part of Rejection criteria. All others will be considered Errors.”

Variable Name (Codelist Name)

TSVAL when TSPARMCD value is SEXPOP (SEXPOP)

TSVAL when TSPARMCD value is ADDON (NY)

TSVAL when TSPARMCD value is RANDOM (NY)

The following table can help us better understand this validation rule. Take the SEXPOP parameter from the TS domain for example, for a specific clinical trial, the participants can only be either female only, male only, or both genders, it can only be one of these three submission values as listed below, it can't be anything else, otherwise, a reject message will be generated during the compliance check.

DATASET	VARIABLE	CONDITIONAL_VALUE1	CTLIST_DESCRIPTION	SUBMISSION_VALUE	DECODE
TS	TSVAL	SEXPOP	Sex of Participants	BOTH	
TS	TSVAL	SEXPOP	Sex of Participants	F	Female
TS	TSVAL	SEXPOP	Sex of Participants	M	Male

Reject message #3: NULL value in variable marked as Required

This is the validation rule of FDAC018, PMDA SD0002, and it applies to all SDTM domains. The description for this validation rule is: “Required variables (where Core attribute is 'Req') cannot be NULL for any records.”

PMDA also noted the following exceptions: “Violation of SD0002 will be considered part of Rejection criteria, except for the following variables in DM domain that will be considered Errors.”

Variable Name

ARMCD

ARM

ACTARMCD

ACTARM

In this validation rule, PMDA specifies that except the above listed four treatment arm related variables in Demographics (DM) domain, for all other required variables per SDTM IG, their value can't be null for any records. Take the Subject Elements (SE) domain for example, STUDYID, DOMAIN, USUBJID, SESEQ, ETCD, SESTDTC are all required variables, their values can't be null for any records.

Reject message #4: SDTM required variable not found

This is the validation rule of FDAC017, PMDA SD0056, and it applies to all SDTM domains. The description for this validation rule is: "Variables described in SDTM IG as Required must be included in the dataset." For example, variable USUBJID is required in most of the SDTM domains, you can't skip this variable in these datasets.

Reject message #5: Incompatible data source

This is the validation rule of FDAC013, PMDA SD0062, and it applies to all SDTM domains. The description for this validation rule is: "Domain table must have a valid format (e.g. SAS transport (XPORT) v.5 or text-delimited)."

Study teams generally use the .sas7bdat format for the SDTM datasets, however, as listed above, both FDA and PMDA require the format to be either XPORT file or text-delimited file. Therefore, before you load your data into the P21E compliance check tool, make sure to transform the data into the right format first.

Reject message #6: Subject is not present in DM domain

This is the validation rule of FDAC040, PMDA SD0064, and it applies to all SDTM domains. The description for this validation rule is: "All Subjects (USUBJID) must be present in Demographics (DM) domain."

The DM includes a set of essential standard variables that describe each subject in a clinical study. It is the parent domain for all other observations for human clinical subjects, and it must have the complete list of all subjects in the clinical trial. However, quite often, due to one data issue or another, we do see patients that don't belong to the current study/trial are present in the raw datasets, especially in the lab related datasets. In this kind of situation, the study team needs to communicate with data management group to remove these patients from raw data.

Reject message #7: Missing DM dataset

This is the validation rule of FDAC001, PMDA SD1020, and it only applies to DM domain. The description for this validation rule is: "Demographics (DM) dataset must be included in every submission."

Without the DM domain, it will not only generate the reject message per PMDA validation rule as listed above, it will also trigger FDA's technical rejection since one of its technical rejection criteria is "The demographic datasets (DM) and Define.xml must be submitted for SEND and SDTM, and the subject level analysis dataset (ADSL) and Define.xml must be submitted for ADaM."

Reject message #8: Variable prohibited for use in SDTM

This is the validation rule of FDAC028, PMDA SD1073, and it applies to all SDTM domains. The description for this validation rule is: "Variables described in IG as inappropriate for usage must be not included in the dataset."

For example, in the Subject Elements (SE) domain, it would be inappropriate to add the variables that support time points (--TPT, --TPTNUM, --ELTM, --TPTREF, and --RFTDTC), since the topic of this dataset is Elements; same for the Subject Visits (SV) domain, it is also inappropriate to include the above mentioned the time points variables since the topic of this data is visits.

Reject message #9: Variable which can be used only in SEND

This is the validation rule of FDAC074, PMDA SD1074, and it applies to all SDTM domains. The description for this validation rule is: "Variables designed only for SEND pre-clinical studies must be not included in the SDTM dataset."

The Standard for Exchange of Nonclinical Data (SEND) is an implementation of the SDTM standard for nonclinical studies. SEND specifies a way to collect and present nonclinical data in a consistent format. There are some variables that were designed for SEND, but not for SDTM human clinical trials. In section 2.7 of the SDTM IG version 3.3, it stated that the following variables are defined for use in the SEND, and they should NEVER be used in the submission of SDTM-based data for human clinical trials:

- --USCHFL (Interventions, Events, Findings)
- --DTHREL (Findings)
- --EXCLFL (Findings)
- --REASEX (Findings)
- --IMPLBL (Findings)
- FETUSID (Identifiers)
- --DETECT (Timing Variables)
- --NOMDY (Timing Variables)
- --NOMLBL (Timing Variables)

The following variables can be used for non-clinical studies (SEND) but must NEVER be used in the Demographics domain for human clinical trials, where all subjects are human.

- SPECIES (Demographics)
- STRAIN (Demographics)
- SBSTRAIN (Demographics)

CONCLUSION

This paper presented an in-depth review of all the nine SDTM reject validation rules per PMDA severity level. It will help the SDTM datasets specification authors and programmers to stay ahead of the game and to avoid the potential pitfalls down the road of SDTM datasets creation, compliance check and submission. Knowledge of these rules can help the study team to build a solid foundation for a successful submission and approval from all of the regulatory agencies around the world.

REFERENCES

CDISC Standard for Exchange of Nonclinical Data Implementation Guide (Version 3.1)

<http://cdisc.org/send>

Study Data Tabulation Model Implementation Guide (Version 3.3)

<https://wiki.cdisc.org/display/SDTMIG33/SDTMIG+v3.3>

Study Data Tabulation Model (SDTM) V1.3

<https://www.cdisc.org/standards/foundational/sdtm>

Pinnacle 21 Enterprise validation rules

<https://www.pinnacle21.com/validation-rules/sdtm>

Technical Rejection Criteria for Study Data

<https://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/FormsSubmissionRequirements/ElectronicSubmissions/UCM523539.pdf>

ACKNOWLEDGMENTS

Many thanks to Yurong Dai for her timely feedback, inputs and review.

CONTACT INFORMATION

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

Lixiang Larry Liu

liu_lixiang@lilly.com

Yuqin Helen Li

li_yuqin_helen@lilly.com